

# Random walks on Baumslag–Solitar groups

Johannes Cuno\*

Graz University of Technology

cuno@math.tugraz.at

Ecaterina Sava-Huss†

Cornell University

sava-huss@cornell.edu

## Abstract

The aim of this paper is to give a description of the Poisson–Furstenberg boundary of random walks on non-amenable Baumslag–Solitar groups. After a short introduction to Baumslag–Solitar groups and their geometry, we change our focus to random walks on these groups. The Poisson–Furstenberg boundary is a probabilistic model for the long-time behaviour of random walks. For random walks on non-amenable Baumslag–Solitar groups we identify the Poisson–Furstenberg boundary in terms of the boundary of the hyperbolic plane and the space of ends of the associated Bass–Serre tree using Kaimanovich’s strip criterion. The precise statement can be found in Theorem 5.11 on page 25.

**Keywords:** random walk, Poisson–Furstenberg boundary, strip criterion, Baumslag–Solitar group, compactification.

**MSC 2010 classes:** 60J50 (primary), 60G50, 20F65, 50C63.

## 1 Introduction

For any two non-zero integers  $p$  and  $q$  the Baumslag–Solitar group  $BS(p, q)$  is given by the presentation  $BS(p, q) = \langle a, b \mid ab^p a^{-1} = b^q \rangle$ . These groups were introduced by Baumslag and Solitar in [BS62], who identified  $BS(2, 3)$  as the first example of a two-generator one-relator non-Hopfian group and thus answered a question by B. H. Neumann, see [Neu54]. Later on, it was shown that  $BS(p, q)$  is Hopfian if and only if  $|p| = 1$  or  $|q| = 1$  or  $\mathcal{P}(p) = \mathcal{P}(q)$ , where  $\mathcal{P}(x)$  denotes the set of prime divisors of  $x$ , see [BS62] and [Mes72].

After reviewing some fundamental properties of Baumslag–Solitar groups, we shall consider random walks on these groups. Such a random walk is constructed as follows. First, we choose a probability measure  $\mu$  on  $BS(p, q)$  such that the support of  $\mu$  generates  $BS(p, q)$  as a semigroup. Then, the random walk starts at the identity element and proceeds with independent  $\mu$ -distributed increments each of which is multiplied from the right to the current state.

---

\*The first author acknowledges the support of the Austrian Science Fund (FWF): W1230-N13 and P24028-N18.

†The second author acknowledges the support of the Austrian Science Fund (FWF): J3575-N26.

The Poisson–Furstenberg boundary was introduced by Furstenberg in [Fur63] and [Fur71]. It is a probabilistic model for the long-time behaviour of the random walk, and simultaneously provides a way to represent all bounded harmonic functions on the state space. In the early 1990s, Kaimanovich considered random walks on  $BS(1,2)$ . Under the assumption of finite first moment, he identified their Poisson–Furstenberg boundary geometrically, see [Kai91, Theorem 5.1]. In particular, he showed that the latter is trivial if the random walk has no vertical drift.

For random walks on non-amenable groups the situation is different because their Poisson–Furstenberg boundary can never be trivial. This motivates the present paper, in which we study random walks on non-amenable Baumslag–Solitar groups and their Poisson–Furstenberg boundary.

The paper is organised as follows. In Section 2, we discuss some algebraic and geometric properties of Baumslag–Solitar groups  $BS(p,q)$  with  $1 \leq p < q$ . We explain how these groups can be understood through their natural projections to the Bass–Serre tree  $\mathbb{T}$  and the hyperbolic plane  $\mathbb{H}$ . Afterwards, we recall the construction of the space of ends  $\partial\mathbb{T}$  and the boundary  $\partial\mathbb{H}$ . They shall later be used to associate a geometric boundary to  $BS(p,q)$ . In Sections 3 and 4, we discuss random walks on countable groups and define the notion of Poisson–Furstenberg boundary. We outline a few classical results and state Kaimanovich’s strip criterion, which is an important tool to identify the Poisson–Furstenberg boundary geometrically. In Section 5, we consider random walks on  $BS(p,q)$  with  $1 \leq p < q$ . In order to ensure that the natural projections of the random walk to  $\mathbb{H}$  and  $\mathbb{T}$  converge almost surely to random elements in  $\partial\mathbb{H}$  and  $\partial\mathbb{T}$  respectively, we need to make suitable assumptions on the moments. If the random walk has vertical drift, we need to assume finite first moment. Otherwise, the situation is much more subtle, and we need to assume finite  $(2+\varepsilon)$ -th moment. The fact that the projections converge almost surely allows us to endow the Cartesian product  $\partial\mathbb{H} \times \partial\mathbb{T}$ , or occasionally just its second component  $\partial\mathbb{T}$ , with the Borel  $\sigma$ -algebra and a hitting measure. Finally, Kaimanovich’s strip criterion shows that the resulting probability space is isomorphic to the Poisson–Furstenberg boundary.

For the part of the paper up to Lemma 5.6, we will assume that the two non-zero integers  $p$  and  $q$  satisfy  $1 \leq p < q$ . In Lemma 5.6 and in the subsequent results, we have to restrict ourselves to the non-amenable subcase  $1 < p < q$ . The main result is Theorem 5.11 on page 25. In the Appendix, we will explain how to obtain similar results for the remaining non-amenable cases.

## Acknowledgements

We would like to thank Wolfgang Woess. He suggested this problem and supported us with references and ideas while the research was carried out. Moreover, we are grateful to Vadim Kaimanovich, who made a couple of comments that led to a substantial improvement of the results.

## 2 Baumslag–Solitar groups

### 2.1 Amenability of Baumslag–Solitar groups

The structure of Baumslag–Solitar groups can be studied by means of HNN extensions. Indeed,  $BS(p,q)$  is precisely the HNN extension  $\mathbb{Z} *_\varphi$  with isomorphism  $\varphi : p\mathbb{Z} \rightarrow q\mathbb{Z}$  given by  $\varphi(p) := q$ . This fact allows

us to use the respective machinery, such as Britton's Lemma, see [Bri63], which implies that a freely reduced non-empty word  $w$  over the letters  $a$  and  $b$  and their formal inverses can only represent the identity element  $1 \in \text{BS}(p, q)$  if it contains  $ab^r a^{-1}$  with  $p \mid r$  or  $a^{-1}b^r a$  with  $q \mid r$  as a subword.

Now, one can easily conclude that, if neither  $|p| = 1$  nor  $|q| = 1$ , the elements  $x := a$  and  $y := bab^{-1}$  generate a non-abelian free subgroup. So,  $\text{BS}(p, q)$  is non-amenable. On the other hand, if  $|p| = 1$  or  $|q| = 1$ , a simple calculation shows that the normal subgroup  $\langle\langle b \rangle\rangle \trianglelefteq \text{BS}(p, q)$  is abelian with quotient isomorphic to  $\mathbb{Z}$ . In this case,  $\text{BS}(p, q)$  is solvable and therefore amenable.

As we will discuss in Section 4.3, the distinction between the two cases is of importance when working with random walks.

## 2.2 Projection to the Bass-Serre tree

Assume first that  $1 \leq p < q$ . The Cayley graph  $\mathbb{G}$  of the group  $G := \text{BS}(p, q)$  with respect to the standard generators  $a$  and  $b$  is the directed multigraph with vertex set  $G$ , edge set  $G \times \{a, b\}$ , source function  $s : G \times \{a, b\} \rightarrow G$  given by  $s(g, x) := g$ , and target function  $t : G \times \{a, b\} \rightarrow G$  given by  $t(g, x) := gx$ . Recall that a graph is just a pair consisting of a vertex set and an edge set with the property that every edge is a two-element subset of the vertex set. Every directed multigraph can be converted into a graph by ignoring the direction and the multiplicity of the edges and deleting the loops. For the purpose of this paper it is sufficient to think of  $\mathbb{G}$  as a graph, and we shall tacitly do so.

Consider the illustration of  $\mathbb{G}$  in Figure 1. Intuitively speaking, we may look at it from the side to see the associated Bass-Serre tree. Formally, let  $B := \langle b \rangle \leq G$  and let  $\mathbb{T}$  be the graph with vertex set  $G/B = \{gB \mid g \in G\}$  and edge set  $\{\{gB, gaB\} \mid g \in G\}$ . This graph is actually a tree; it is obviously connected and, by Britton's Lemma, it does not contain any cycle. Note that the canonical projection  $\pi_{\mathbb{T}} : G \rightarrow G/B$  given by  $\pi_{\mathbb{T}}(g) := gB$  is a weak graph homomorphism from  $\mathbb{G}$  to  $\mathbb{T}$ , i.e. whenever the vertices  $g$  and  $h$  are adjacent in  $\mathbb{G}$ , their images  $gB$  and  $hB$  either agree or they are adjacent in  $\mathbb{T}$ .

**Remark 2.1 (“levels”)** Consider the infinite cyclic group  $\mathbb{Z}$  and the map  $\lambda : \{a, b\} \rightarrow \mathbb{Z}$  given by  $\lambda(a) := 1$  and  $\lambda(b) := 0$ . The latter can be uniquely extended to a group homomorphism  $\lambda : G \rightarrow \mathbb{Z}$ . Indeed, the equation  $\lambda(a) + p \cdot \lambda(b) - \lambda(a) = q \cdot \lambda(b)$  holds in  $\mathbb{Z}$  so that we can apply von Dyck's Theorem to extend  $\lambda$ , see e.g. [Rot95, p. 346, fn. 2]. Since  $\lambda(b) = 0$ , the group homomorphism  $\lambda : G \rightarrow \mathbb{Z}$  is constant on the cosets from  $G/B$  and therefore induces a well-defined map  $\tilde{\lambda} : G/B \rightarrow \mathbb{Z}$  given by  $\tilde{\lambda}(gB) := \lambda(g)$ . We shall think of  $\lambda$  and  $\tilde{\lambda}$  as level functions, they assign a level to every vertex of  $\mathbb{G}$  and  $\mathbb{T}$ , respectively.

**Lemma 2.2** Every vertex  $gB$  of  $\mathbb{T}$  has exactly  $p + q$  neighbours;  $p$  of them are one level below and  $q$  of them are one level above the vertex.

*Proof.* By construction, the levels of two adjacent vertices always differ exactly by 1. The defining relation  $ab^p a^{-1} = b^q$  (“ $\Leftarrow$ ”) and Britton's Lemma (“ $\Rightarrow$ ”) imply that  $gaB = gb^r aB$  if and only if  $q \mid r$ , whence the vertex  $gB$  has exactly  $q$  neighbours above. Similarly, it has exactly  $p$  neighbours below because  $ga^{-1}B = gb^r a^{-1}B$  if and only if  $p \mid r$ .  $\square$

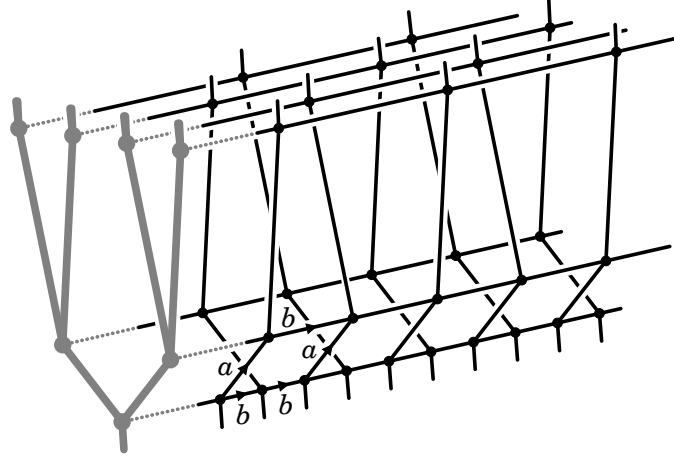


Figure 1: The Cayley graph  $\mathbb{G}$  of  $BS(1,2)$  with respect to the standard generators  $a$  and  $b$ .

### 2.3 Projection to the hyperbolic plane

The second projection captures the information that is obtained by looking at  $\mathbb{G}$  from the front. In order to construct it, we introduce another group. Let  $\text{Aff}^+(\mathbb{R})$  be the set of all affine transformations of the real line that preserve the orientation, i. e. all maps  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  of the form  $\varphi(x) = \alpha x + \beta$  with  $\alpha, \beta \in \mathbb{R}$  and  $\alpha > 0$ . This set endowed with the composition  $(\varphi_2 \circ \varphi_1)(x) := (\alpha_2 \alpha_1)x + (\alpha_2 \beta_1 + \beta_2)$  forms a group. As in the construction of the level function  $\lambda : G \rightarrow \mathbb{Z}$  in Remark 2.1, consider the map  $\pi_{\text{Aff}^+(\mathbb{R})} : \{a, b\} \rightarrow \text{Aff}^+(\mathbb{R})$  given by  $\pi_{\text{Aff}^+(\mathbb{R})}(a) := (x \mapsto \frac{q}{p} \cdot x)$  and  $\pi_{\text{Aff}^+(\mathbb{R})}(b) := (x \mapsto x + 1)$ . Due to von Dyck's Theorem, it can be uniquely extended to a group homomorphism  $\pi_{\text{Aff}^+(\mathbb{R})} : G \rightarrow \text{Aff}^+(\mathbb{R})$ . The group  $\text{Aff}^+(\mathbb{R})$  has a geometric interpretation. In order to describe it, let  $\mathbb{H}$  be the hyperbolic plane as per the Poincaré half-plane model, i. e.  $\mathbb{H} = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$ , endowed with the standard metric

$$d_{\mathbb{H}}(z_1, z_2) := \ln \left( \frac{|z_1 - \bar{z}_2| + |z_1 - z_2|}{|z_1 - \bar{z}_2| - |z_1 - z_2|} \right) = \text{arcosh} \left( 1 + \frac{|z_1 - z_2|^2}{2\text{Im}(z_1)\text{Im}(z_2)} \right).$$

The elements of  $\text{Aff}^+(\mathbb{R})$  can be thought of as isometries of the hyperbolic plane  $\mathbb{H}$ , which are precisely the maps  $\varphi : \mathbb{H} \rightarrow \mathbb{H}$  of the form

$$\varphi(z) = \frac{\alpha z + \beta}{\gamma z + \delta} \quad \text{or} \quad \varphi(z) = \frac{\alpha \cdot (-\bar{z}) + \beta}{\gamma \cdot (-\bar{z}) + \delta} \quad \text{with} \quad \alpha, \beta, \gamma, \delta \in \mathbb{R} \text{ and } \alpha\delta - \beta\gamma > 0,$$

see e. g. [Bea83, Theorem 7.4.1].<sup>1</sup> Now, we are ready to construct the second projection  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$ . Pick an element  $g \in G$ , map it via  $\pi_{\text{Aff}^+(\mathbb{R})}$  to  $\text{Aff}^+(\mathbb{R})$ , think of the latter as an isometry of  $\mathbb{H}$ , and evaluate it at  $i \in \mathbb{H}$ . The following lemma illustrates this construction.

**Lemma 2.3** *For every  $g \in G$  the point  $\pi_{\mathbb{H}}(ga) \in \mathbb{H}$  is above the point  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ ; the two points have the same real part and their distance is  $\ell_a := \ln(\frac{q}{p})$ . Similarly, for every  $g \in G$  the point  $\pi_{\mathbb{H}}(gb) \in \mathbb{H}$  is right from the point  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ ; the two points have the same imaginary part and their distance is  $\ell_b := \ln(\frac{3+\sqrt{5}}{2})$ . So, in some way, we are actually looking at  $\mathbb{G}$  from the front.*

<sup>1</sup>To be more precise, the elements of  $\text{Aff}^+(\mathbb{R})$  correspond to the orientation-preserving isometries of  $\mathbb{H}$  that fix  $\infty \in \partial\mathbb{H}$ , which is defined in Section 2.5. The orientation-reversing isometries of  $\mathbb{H}$  that fix  $\infty \in \partial\mathbb{H}$  will be crucial for the investigation of Baumslag-Solitar groups  $BS(p, q)$  with  $1 < p < -q$  in the Appendix.

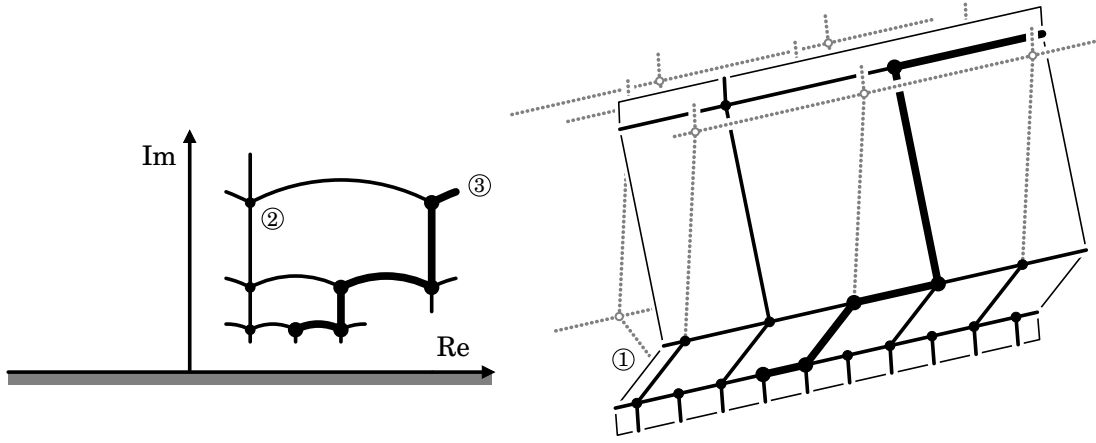


Figure 2: A part of a discrete hyperbolic plane  $\mathbb{G}_v$  (right) and its projection to  $\mathbb{H}$  (left).

*Proof.* This is clear for  $g = 1$ . Now, pick an arbitrary element  $g \in G$ . The points  $\pi_{\mathbb{H}}(ga) \in \mathbb{H}$  and  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$  are obtained by applying  $\pi_{\text{Aff}^+(\mathbb{R})}(g)$  to the points  $\pi_{\mathbb{H}}(a) \in \mathbb{H}$  and  $\pi_{\mathbb{H}}(1) \in \mathbb{H}$ .<sup>2</sup> But since  $\pi_{\text{Aff}^+(\mathbb{R})}(g)$  is the composition of a dilation  $z \mapsto \alpha z$  and a translation  $z \mapsto z + \beta$ , the relative position of the two points is preserved. The same argument works for the second assertion, which completes the proof.  $\square$

## 2.4 Discrete hyperbolic plane

Here and throughout the paper, we use the symbol  $\mathbb{N}_0$  to denote the non-negative integers and the symbol  $\mathbb{N}$  to denote the strictly positive integers.

**Definition 2.4 (“path”, “reduced path”)** *Given a graph with vertex set  $V$ , we consider finite paths  $v : \{0, 1, \dots, n\} \rightarrow V$ , infinite paths  $v : \mathbb{N}_0 \rightarrow V$ , and doubly infinite paths  $v : \mathbb{Z} \rightarrow V$ . In any case, being a path means that for every possible choice of  $k$  the vertices  $v(k)$  and  $v(k + 1)$  are adjacent in the graph. Moreover, we say that a path is reduced if for every possible choice of  $k$  the vertices  $v(k)$  and  $v(k + 2)$  are distinct.*

Fix an ascending doubly infinite path  $v : \mathbb{Z} \rightarrow G/B$  in the tree  $\mathbb{T}$ . Ascending refers to the level function defined in Remark 2.1, and it means that for every  $k \in \mathbb{Z}$  the vertex  $v(k)$  is located above the preceding vertex  $v(k - 1)$ . Now, let  $G_v$  be the full  $\pi_{\mathbb{T}}$ -preimage of  $v(\mathbb{Z})$ , i. e. the set consisting of all  $g \in G$  such that the image  $\pi_{\mathbb{T}}(g)$  is traversed by  $v$ . The subgraph  $\mathbb{G}_v \leq \mathbb{G}$  spanned by  $G_v$ , see ① in Figure 2, is obviously connected so that the graph distance  $d_{\mathbb{G}_v}$  becomes a metric. This subgraph is sometimes referred to as discrete hyperbolic plane or plane of bricks, which makes particular sense in light of Proposition 2.5. Variations of the latter have already been used in the literature, e. g. in [Anc88], [FM98], and [CFM04]. Concerning [Anc88], see also the remark in [CW92, p. 382].

**Proposition 2.5** *The restriction  $\pi_{\mathbb{H}}|_{G_v} : G_v \rightarrow \mathbb{H}$  is a quasi-isometry, even a bi-Lipschitz map, between the graph  $\mathbb{G}_v$  endowed with the graph distance  $d_{\mathbb{G}_v}$  and the hyperbolic plane  $\mathbb{H}$  endowed with the*

<sup>2</sup>Note that the equation  $(\varphi_2 \circ \varphi_1)(x) = \varphi_2(\varphi_1(x))$  remains true when replacing  $x \in \mathbb{R}$  by  $z \in \mathbb{H}$ . Therefore, we may actually conclude that  $\pi_{\mathbb{H}}(ga) = \pi_{\text{Aff}^+(\mathbb{R})}(ga)(i) = (\pi_{\text{Aff}^+(\mathbb{R})}(g) \circ \pi_{\text{Aff}^+(\mathbb{R})}(a))(i) = \pi_{\text{Aff}^+(\mathbb{R})}(g)(\pi_{\text{Aff}^+(\mathbb{R})}(a)(i)) = \pi_{\text{Aff}^+(\mathbb{R})}(g)(\pi_{\mathbb{H}}(a))$ .

standard metric  $d_{\mathbb{H}}$ .

*Proof.* We realise the edges of the graph  $\mathbb{G}_v$  geometrically. Whenever two vertices  $g, h \in G_v$  are adjacent, we connect their images  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$  and  $\pi_{\mathbb{H}}(h) \in \mathbb{H}$  by a geodesic in  $\mathbb{H}$ . In order to avoid confusion, we refer to these images as  $\mathbb{H}$ -vertices and to the geodesics between them as  $\mathbb{H}$ -edges. By the proof of Lemma 2.2 and by Lemma 2.3, the  $\mathbb{H}$ -vertices and  $\mathbb{H}$ -edges yield a tessellation of the hyperbolic plane with isometric bricks of the following shape. The  $\mathbb{H}$ -vertices of each brick are located on two distinct horizontal lines; on the upper one there are  $p+1$  and on the lower one there are  $q+1$ . In either case, the  $\mathbb{H}$ -vertices are connected by  $\mathbb{H}$ -edges of length  $\ell_b$  to form a chain (= piecewise geodesic curve). Due to the curvature, both the two leftmost and the two rightmost  $\mathbb{H}$ -vertices are located precisely above each other and connected by vertical  $\mathbb{H}$ -edges of length  $\ell_a$ , see ② in Figure 2 and Figure 3. Since the bricks are uniformly bounded and cover the hyperbolic plane  $\mathbb{H}$ , the restriction  $\pi_{\mathbb{H}}|_{G_v} : G_v \rightarrow \mathbb{H}$  is certainly quasi-surjective.

Pick any two vertices  $g, h \in G_v$ . We aim to estimate the distances  $d_{\mathbb{G}_v}(g, h)$  and  $d_{\mathbb{H}}(\pi_{\mathbb{H}}(g), \pi_{\mathbb{H}}(h))$  by multiples of each other. First, choose a path of minimal length from  $g$  to  $h$  in  $\mathbb{G}_v$ . It corresponds to a chain of  $\mathbb{H}$ -edges from  $\pi_{\mathbb{H}}(g)$  to  $\pi_{\mathbb{H}}(h)$ , see ③ in Figure 2. This chain consists of  $d_{\mathbb{G}_v}(g, h)$  many  $\mathbb{H}$ -edges, each of which has length at most  $\max\{\ell_a, \ell_b\}$ . Hence,

$$d_{\mathbb{H}}(\pi_{\mathbb{H}}(g), \pi_{\mathbb{H}}(h)) \leq d_{\mathbb{G}_v}(g, h) \cdot \max\{\ell_a, \ell_b\}.$$

For the converse estimate, let us make the following auxiliary definition. Every point  $x \in \mathbb{H}$  that is not in the interior of a brick is either an  $\mathbb{H}$ -vertex, in which case we define  $x'$  to be  $x$ , or it is in the interior of an  $\mathbb{H}$ -edge, in which case we define  $x'$  to be one of the endpoints of the  $\mathbb{H}$ -edge, whichever is closer. In the case that  $x$  is precisely in the middle of the  $\mathbb{H}$ -edge, we choose the left endpoint rather than the right one and the lower endpoint rather than the upper one. With this notion in mind, consider the geodesic  $\gamma$  from  $\pi_{\mathbb{H}}(g)$  to  $\pi_{\mathbb{H}}(h)$ , see Figure 3. Whenever  $\gamma$  traverses the interior of a brick  $B$ , it enters the interior at some point  $x \in \partial B$  and leaves it at some other point  $y \in \partial B$ . In this situation, approximate the part of  $\gamma$  from  $x$  to  $y$  by a chain of  $\mathbb{H}$ -edges from  $x'$  to  $y'$ . We may choose this chain such that, whenever  $x' = y'$ , the chain has no  $\mathbb{H}$ -edge at all and, otherwise, the number of  $\mathbb{H}$ -edges in the chain is at most  $c := \lfloor \frac{1}{2} \cdot (p+q+2) \rfloor$ . But, by a compactness argument, there is an  $\varepsilon > 0$  such that if the part of  $\gamma$  has length smaller than  $\varepsilon$ , then  $x' = y'$  and the chain has no  $\mathbb{H}$ -edge at all. Therefore, we may conclude that

$$\text{number of } \mathbb{H}\text{-edges in the chain} \leq \frac{c}{\varepsilon} \cdot \text{length of the part of } \gamma.$$

It is not hard to see that if we do this for every brick  $B$  whose interior is traversed by  $\gamma$ , we finally obtain a chain of  $\mathbb{H}$ -edges from  $\pi_{\mathbb{H}}(g)$  to  $\pi_{\mathbb{H}}(h)$ . Depending on whether a part of  $\gamma$  originally traversed the interior of a brick or ran along an  $\mathbb{H}$ -edge, we may estimate the number of  $\mathbb{H}$ -edges approximating it by  $\frac{c}{\varepsilon}$ , by  $\frac{1}{\ell_a}$ , or by  $\frac{1}{\ell_b}$  times its length. Hence,

$$d_{\mathbb{G}_v}(g, h) \leq d_{\mathbb{H}}(\pi_{\mathbb{H}}(g), \pi_{\mathbb{H}}(h)) \cdot \max \left\{ \frac{c}{\varepsilon}, \frac{1}{\ell_a}, \frac{1}{\ell_b} \right\}.$$

□

**Remark 2.6** Note that the horizontal lines mentioned in the proof of Proposition 2.5 are horospheres, and by no means geodesics. For example, one may pick such a horizontal line and observe that the part of the line contained in the closed disc  $D \subseteq \mathbb{H}$  with centre  $i \in \mathbb{H}$  and radius  $n \in \mathbb{N}$  has a length growing exponentially in  $n$ , see also Figure 11 on page 28.

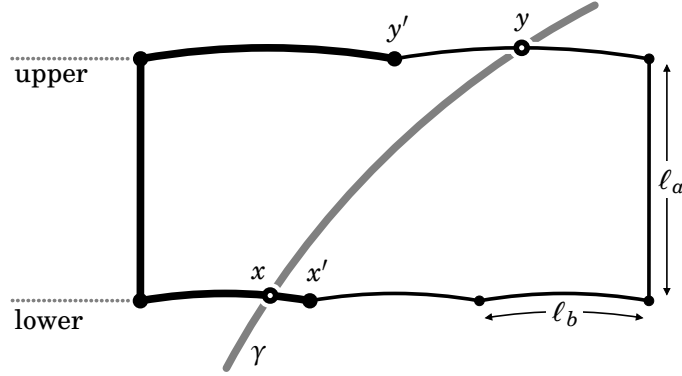


Figure 3: Approximation of the geodesic  $\gamma$  in the case  $p = 2$  and  $q = 3$ .

**Remark 2.7** The level of a vertex  $g \in G$  can be recovered both from  $\pi_{\mathbb{T}}(g) \in G/B$  and from  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ . In fact, the image of  $\pi_{\mathbb{T}} \times \pi_{\mathbb{H}} : G \rightarrow G/B \times \mathbb{H}$  is contained in the horocyclic product of the tree  $\mathbb{T}$  and the hyperbolic plane  $\mathbb{H}$ , which is sometimes referred to as treebolic space, see [BSSW15] for details.

## 2.5 Compactifications

Both the tree  $\mathbb{T}$  and the hyperbolic plane  $\mathbb{H}$  have a natural compactification. In case of  $\mathbb{T}$ , it is the end compactification, which can be constructed as follows. Fix a base point, say  $B \in G/B$ , and consider the set  $\hat{\mathbb{T}}$  of all reduced paths that start in  $B$ , be they finite or infinite. The endpoint map yields a one-to-one correspondence between the finite paths and the vertices  $G/B$ . We may therefore think of  $G/B$  as a subset of  $\hat{\mathbb{T}}$ . The set  $\hat{\mathbb{T}}$  can be endowed with the metric

$$d_{\hat{\mathbb{T}}}(x, y) = \begin{cases} 2^{-|x \wedge y|} & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}.$$

Here, the symbol  $|x \wedge y|$  denotes the number of edges the two paths run together until they separate, i. e.  $|x \wedge y| = \max\{k \in \mathbb{N}_0 \mid x(k) \text{ and } y(k) \text{ are both defined and } x(k) = y(k)\}$ , see ① in Figure 4. Hence, the later the paths separate the closer they are. The set  $\hat{\mathbb{T}}$  endowed with the metric  $d_{\hat{\mathbb{T}}}$  is a compact metric space that contains  $G/B$  as a discrete and dense subset. The complement of  $G/B$  is the set of infinite paths, it is usually denoted by  $\partial\mathbb{T}$  and called the space of ends.

In case of  $\mathbb{H}$ , we temporarily switch to the Poincaré disc model. More precisely, instead of working in the half-plane  $\mathbb{H} = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$ , we consider the open unit disc  $\mathbb{D} := \{z \in \mathbb{C} \mid |z| < 1\}$ . The Cayley transform  $W : \mathbb{H} \rightarrow \mathbb{D}$  given by  $W(z) := \frac{z-i}{z+i}$  is one possibility to convert between the two models. Since we are currently interested in the topological structure, let us highlight that the hyperbolic topology on  $\mathbb{D}$  is the one induced by the Cayley transform, i. e. the one that turns the Cayley transform into a homeomorphism. It happens to agree with the standard topology on  $\mathbb{D}$ . So, topologically speaking, the hyperbolic plane in the Poincaré disc model is just a subspace of the complex plane  $\mathbb{C}$ . We may therefore compactify it by taking the closed unit disc  $\hat{\mathbb{D}} := \{z \in \mathbb{C} \mid |z| \leq 1\}$ , see Figure 4. In order to translate this compactification back to the Poincaré half-plane model, we first extend both the domain and the codomain of the Cayley transform so that we obtain a bijection  $W : \mathbb{H} \cup \mathbb{R} \cup \{\infty\} \rightarrow \hat{\mathbb{D}}$ , and then apply its inverse. The resulting space  $\hat{\mathbb{H}} := \mathbb{H} \cup \mathbb{R} \cup \{\infty\}$  is our compactification. It is, once again, endowed with



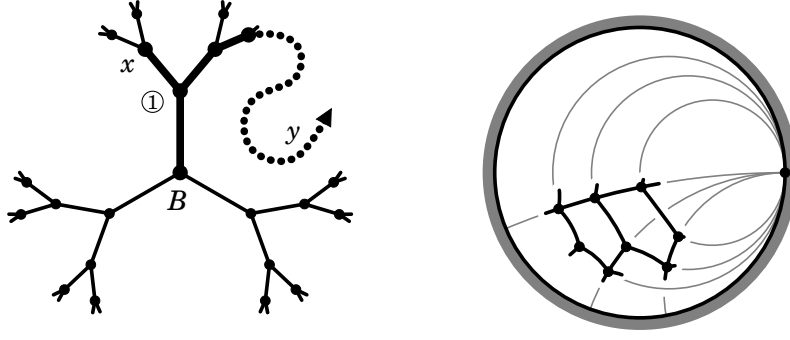


Figure 4: The space of ends (left) and the hyperbolic boundary in the Poincaré disc model (right).

the induced topology, and thus a compact space that contains  $\mathbb{H}$  as a dense subset. The complement of  $\mathbb{H}$  is the union  $\mathbb{R} \cup \{\infty\}$ , it is usually denoted by  $\partial\mathbb{H}$  and called the hyperbolic boundary. Having introduced the hyperbolic boundary this way, the following lemma gives us a helpful criterion for convergence. Its proof is elementary and we leave it to the reader.

**Lemma 2.8** *A sequence  $(x_0, x_1, \dots)$  in  $\mathbb{H}$  converges to  $\infty \in \partial\mathbb{H}$  if and only if the absolute values  $|x_n|$  tend to infinity. Moreover, it converges to a point  $r \in \partial\mathbb{H} \setminus \{\infty\}$  if and only if it does with respect to the standard topology on the complex plane  $\mathbb{C}$ .*

### 3 Random walks on groups

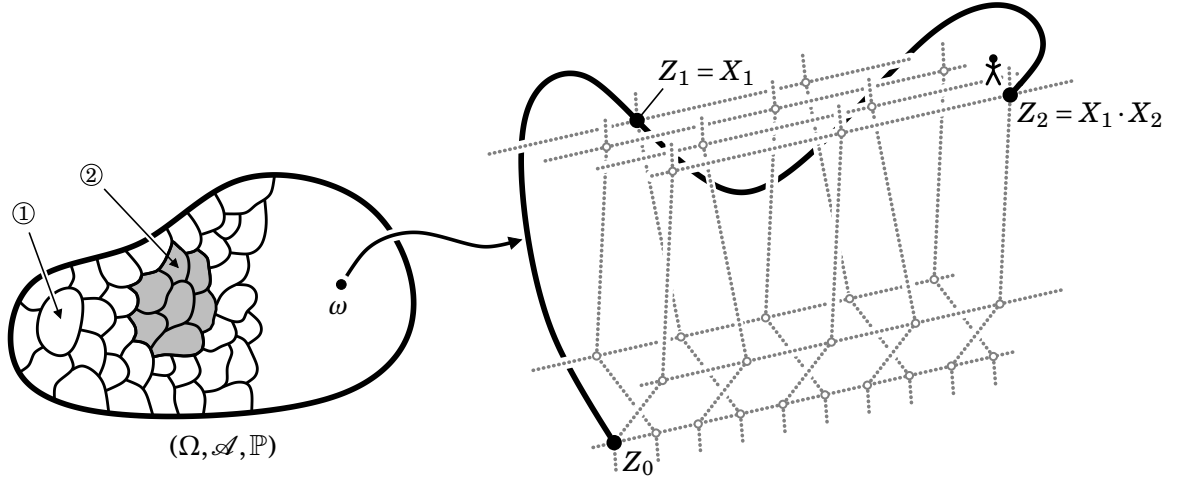
#### 3.1 Basic notions

The aim of the current work is to study random walks on Baumslag–Solitar groups. Before doing so, we fix the notation. Given a countable state space  $X$ , an initial probability measure  $\vartheta : X \rightarrow [0, 1]$ , and transition probabilities  $p : X \times X \rightarrow [0, 1]$ , we are interested in the Markov chain  $Z = (Z_0, Z_1, \dots)$  that starts according to  $\vartheta$  and proceeds according to  $p$ . Formally, we construct the probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ , where  $\Omega := \{(x_0, x_1, \dots) \mid \forall n \in \mathbb{N}_0 : x_n \in X\}$  is the set of trajectories,  $\mathcal{A}$  is the product  $\sigma$ -algebra, and  $\mathbb{P}$  is the probability measure induced by  $\vartheta$  and  $p$ . The projections  $Z_n : \Omega \rightarrow X$  given by  $Z_n(x_0, x_1, \dots) := x_n$  become random variables that constitute the Markov chain. For details on the terminology used above, see e. g. [Kle14, §1], and for a gentle introduction to discrete Markov chains, see e. g. [Woe09, §1]. We will use the term random walk instead of Markov chain.

Next, let us assume that  $X$  is a countable group  $G$ , in which case we may study random walks whose transition probabilities are adapted to the group structure. In order to do so, we first pick a probability measure  $\mu : G \rightarrow [0, 1]$  whose support  $\text{supp}(\mu) = \{g \in G \mid \mu(g) > 0\}$  generates  $G$  as a semigroup, see also Remark 3.1 below. Then, we consider the random walk given by the following data. The initial probability measure  $\vartheta : G \rightarrow [0, 1]$  puts all mass on the identity element  $1 \in G$  and the transition probabilities  $p : G \times G \rightarrow [0, 1]$  are given by  $p(g, h) := \mu(g^{-1}h)$ .

We could also have said  $p(g, gx) := \mu(x)$ , which leads to a handy interpretation. The random walk starts at the identity element and has independent  $\mu$ -distributed increments each of which is multiplied




 Figure 5: The first steps of a random walk on  $BS(1,2)$ .

from the right to the current state. Therefore,  $Z_0 = 1$  a.s. (=almost surely) and for every  $n \in \mathbb{N}$  we may decompose  $Z_n = X_1 \cdot \dots \cdot X_n$ , where  $X_1, X_2, \dots$  is a sequence of independent  $\mu$ -distributed random variables, see the right-hand side of Figure 5.

**Remark 3.1** *Since we assume that  $\text{supp}(\mu)$  generates  $G$  as a semigroup, the random walk is irreducible, i. e. any two states can be reached from each other with positive probability. In particular, the following dichotomy holds. Either every state is recurrent, i. e. the return probability is equal to 1, or every state is transient, i. e. the return probability is smaller than 1. In the latter case, the probability that every finite set of states will eventually be left and the random walk escapes to infinity is equal to 1.*

### 3.2 Finite moments

We need to assume that the probability of huge jumps is sufficiently small. The notion of moments helps us to make this assumption rigorous. Given a probability space, e. g.  $(\Omega, \mathcal{A}, \mathbb{P})$  introduced in Section 3.1, and a real valued random variable  $X : \Omega \rightarrow \mathbb{R}$ , the latter has finite first moment if  $\int |X| d\mathbb{P} < \infty$ . In this case, both  $\int X^+ d\mathbb{P} < \infty$  and  $\int X^- d\mathbb{P} < \infty$ , and we can define the expectation  $\mathbb{E}(X) := \int X^+ d\mathbb{P} - \int X^- d\mathbb{P}$ . Of course, the difference would still make sense if only one of the two integrals was finite. But this is not of relevance for us and when writing  $\mathbb{E}(X)$  we implicitly mean that  $-\infty < \mathbb{E}(X) < \infty$ . More generally, given any non-negative  $k \in \mathbb{R}$ , a real valued random variable  $X : \Omega \rightarrow \mathbb{R}$  has finite  $k$ -th moment if  $\int |X|^k d\mathbb{P} < \infty$ . In our setting, the increments  $X_1, X_2, \dots$  are not real valued random variables but take values in  $G$ , whence we need to specify the hugeness of a jump before talking about finite moments.

**Definition 3.2 (“word metric”, “finite  $k$ -th moment”)** *If  $G$  is a finitely generated group and  $S \subseteq G$  is a finite generating set, then the word metric  $d_S$  on  $G$  is given by*

$$d_S(g, h) := \min\{n \in \mathbb{N}_0 \mid \exists s_1, \dots, s_n \in S : \exists \varepsilon_1, \dots, \varepsilon_n \in \{1, -1\} : g^{-1}h = s_1^{\varepsilon_1} \cdot \dots \cdot s_n^{\varepsilon_n}\}.$$

*Note that the word metric coincides with the distance in the respective Cayley graph. A random variable  $X : \Omega \rightarrow G$  has finite  $k$ -th moment if the image  $d_S(1, X) : \Omega \rightarrow \mathbb{R}$  has finite  $k$ -th moment in the classical sense, i. e. if  $\int d_S(1, X)^k d\mathbb{P} < \infty$ .*

**Remark 3.3** We leave it to the reader to verify that this property does not depend on the choice of the finite generating set  $S \subseteq G$ , see also [Mei08, Lemma 11.37].

### 3.3 Real parts, imaginary parts, and vertical drift

Let us now return to the situation we are interested in, namely that  $G = \text{BS}(p, q)$  with  $1 \leq p < q$ . When working with the projection  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$ , we often consider the imaginary parts  $\text{Im}(\pi_{\mathbb{H}}(g))$  and the real parts  $\text{Re}(\pi_{\mathbb{H}}(g))$  separately, and it is convenient to abbreviate the former by  $A_g$  and the latter by  $B_g$ . Occasionally, we do not need to assume that  $X_1$  has some finite moment but impose this assumption on the images  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$ . The following lemma relates the two situations.

**Lemma 3.4** *If  $X_1$  has finite  $k$ -th moment, then  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$  have finite  $k$ -th moment, too.*

**Remark 3.5** *Before we prove Lemma 3.4, let us note that for every  $g \in G$  the imaginary part  $A_g$  can be expressed in terms of the level  $\lambda(g)$ , namely as  $A_g = \left(\frac{q}{p}\right)^{\lambda(g)}$ . This formula can be shown using either the multiplication in  $\text{Aff}^+(\mathbb{R})$  or Lemma 2.3. Taking the logarithm on both sides yields  $\ln(A_g) = \ln\left(\frac{q}{p}\right) \cdot \lambda(g)$ . So, instead of thinking of  $\ln(A_g)$  we may think of a multiple of  $\lambda(g)$ .*

*Proof of Lemma 3.4.* Let  $S := \{a, b\} \subseteq G$  be the standard generating set. Then

$$\int |\ln(A_{X_1})|^k d\mathbb{P} = \left(\ln\left(\frac{q}{p}\right)\right)^k \cdot \int |\lambda(X_1)|^k d\mathbb{P} \leq \left(\ln\left(\frac{q}{p}\right)\right)^k \cdot \underbrace{\int d_S(1, X_1)^k d\mathbb{P}}_{< \infty} < \infty.$$

Concerning the second assertion, observe that  $d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) \leq \max\{\ell_a, \ell_b\} \cdot d_S(1, g)$ , which can be shown by the same argument as in the proof of Proposition 2.5. This observation allows us to estimate  $\ln(1 + |B_g|)$  by a multiple of  $d_S(1, g)$ . Indeed,

$$\begin{aligned} \ln(1 + |B_g|) &\leq \ln\left(1 + \frac{1}{2} \cdot |B_g|^2 + \sqrt{\left(1 + \frac{1}{2} \cdot |B_g|^2\right)^2 - 1}\right) = \text{arcosh}\left(1 + \frac{1}{2} \cdot |B_g|^2\right) = d_{\mathbb{H}}(i, i + B_g) \\ &\leq d_{\mathbb{H}}(i, A_g \cdot i + B_g) + d_{\mathbb{H}}(A_g \cdot i + B_g, i + B_g) = d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) + |\ln(A_g)| \\ &\leq \max\{\ell_a, \ell_b\} \cdot d_S(1, g) + \ln\left(\frac{q}{p}\right) \cdot |\lambda(g)| \leq \max\{\ell_a, \ell_b\} \cdot d_S(1, g) + \ln\left(\frac{q}{p}\right) \cdot d_S(1, g). \end{aligned}$$

Therefore,

$$\int \ln(1 + |B_{X_1}|)^k d\mathbb{P} \leq \left(\max\{\ell_a, \ell_b\} + \ln\left(\frac{q}{p}\right)\right)^k \cdot \underbrace{\int d_S(1, X_1)^k d\mathbb{P}}_{< \infty} < \infty,$$

which proves the claim.  $\square$

It is easy to construct examples showing that the converse of Lemma 3.4 does not hold. In addition to the moments of  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$ , we will use the notion of vertical drift. Consider a random walk  $Z = (Z_0, Z_1, \dots)$  on  $G$  and its pointwise projection  $\lambda(Z) = (\lambda(Z_0), \lambda(Z_1), \dots)$  to the levels. Since  $\lambda(Z_n) = \lambda(X_1 \cdot \dots \cdot X_n) = \lambda(X_1) + \dots + \lambda(X_n)$ , these projections constitute a random walk on the integers with i. i. d. (=independent and identically distributed) increments.

**Definition 3.6 (“vertical drift”)** *If  $\ln(A_{X_1})$  has finite first moment, then  $\lambda(X_1)$  has finite first moment and we can define the expectation  $\mathbb{E}(\lambda(X_1))$ . The latter is called the vertical drift and denoted by  $\delta$ . We will distinguish between positive vertical drift, i. e.  $\delta > 0$ , negative vertical drift, i. e.  $\delta < 0$ , and no vertical drift, i. e.  $\delta = 0$ , which is the most subtle of the three cases.*

## 4 Poisson–Furstenberg boundary

### 4.1 Lebesgue–Rohlin spaces

The Poisson–Furstenberg boundary is a probabilistic model for the long-time behaviour of a random walk. In order to define it, we need to ensure that we are working with Lebesgue–Rohlin spaces, which are also known as standard probability spaces. For definitions and basic examples we refer to [Roh52], [Hae73], and [Rud90]. Moreover, let us mention the collection of facts in [KKR04, Appendix] and the more informal introduction in [CK12].

The most prominent examples of Lebesgue–Rohlin spaces are discrete probability spaces and the unit interval  $[0, 1]$  endowed with the Lebesgue  $\sigma$ -algebra  $\mathcal{L}$  and the Lebesgue measure  $\lambda$ . In fact, every Lebesgue–Rohlin space is isomorphic<sup>3</sup> either to one of these examples or to the disjoint union of an interval  $[0, \alpha]$  with  $0 < \alpha \leq 1$  and countably many atoms with total mass  $1 - \alpha$ , see [Roh52, §2.4] and [Hae73, Proposition 6].

**Remark 4.1 (“polish spaces”)** *A Polish space is a topological space that is separable, i. e. contains a countable and dense subset, and completely metrisable, i. e. there is a metric that induces the topology and turns the space into a complete metric space. All Polish spaces endowed with the Borel  $\sigma$ -algebra  $\mathcal{B}$  and a Borel measure  $\mu$  become, after completion, examples of Lebesgue–Rohlin spaces, see [Roh52, §2.7] and [Hae73, p. 248, Example 1].*

### 4.2 Poisson–Furstenberg boundary and some of its properties

In light of Remark 4.1, we may observe that the space of trajectories  $\Omega$  introduced in Section 3.1 is the product  $X^{\mathbb{N}_0}$  and can therefore be endowed with the product topology. One can show that the latter is actually a Polish space, see e. g. [Wil70, Theorem 24.11]. Since its Borel  $\sigma$ -algebra agrees with the product  $\sigma$ -algebra  $\mathcal{A}$ , the completion of  $(\Omega, \mathcal{A}, \mathbb{P})$  is a Lebesgue–Rohlin space. From now on, let us assume that, as soon as a measurable space is endowed with a probability measure, we are working with its completion. We may therefore say that  $(\Omega, \mathcal{A}, \mathbb{P})$  is a Lebesgue–Rohlin space.

Since we are interested in the long-time behaviour of the trajectories  $x = (x_0, x_1, \dots) \in \Omega$ , we identify those pairs of trajectories whose tails sooner or later behave identically. More precisely, we define an equivalence relation  $\sim$  on  $\Omega$  by

$$x \sim y : \Longleftrightarrow \exists t_1, t_2 \in \mathbb{N}_0 \forall n \in \mathbb{N}_0 : x_{t_1+n} = y_{t_2+n}.$$

---

<sup>3</sup>We consider probability spaces up to subsets of measure 0. So, we actually mean isomorphic mod 0. Recall that two probability spaces  $(\Omega_1, \mathcal{A}_1, \mathbb{P}_1)$  and  $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$  are isomorphic mod 0 if there are null sets  $N_k \subseteq \Omega_k$  with  $k \in \{1, 2\}$  and a bijection  $\varphi : \Omega_1 \setminus N_1 \rightarrow \Omega_2 \setminus N_2$  which is measurable and measure preserving in both directions.

Note that we allow the times  $t_1$  and  $t_2$  to be different. If we did not, we would end up with the tail boundary instead of the Poisson–Furstenberg boundary. Consider the partition  $\zeta$  of  $\Omega$  into equivalence classes mod  $\sim$ , see ① in Figure 5. This partition induces a sub- $\sigma$ -algebra  $\mathcal{A}_\zeta$  of  $\mathcal{A}$ , consisting of all  $A \in \mathcal{A}$  which are compatible with the partition  $\zeta$ , i. e. which are unions of equivalence classes mod  $\sim$ , see ② in Figure 5. The Poisson–Furstenberg boundary  $(B, \mathcal{B}, \nu)$  is the quotient of  $(\Omega, \mathcal{A}, \mathbb{P})$  with respect to the induced sub- $\sigma$ -algebra  $\mathcal{A}_\zeta$ . More precisely, it is the Lebesgue–Rohlin space  $(\zeta_1, \mathcal{A}_{\zeta_1}, \mathbb{P}|_{\mathcal{A}_{\zeta_1}})$  that consists of the measurable hull  $\zeta_1$  of  $\zeta$ , the induced sub- $\sigma$ -algebra  $\mathcal{A}_{\zeta_1}$ , and the restriction  $\mathbb{P}|_{\mathcal{A}_{\zeta_1}}$  of the probability measure  $\mathbb{P}$  to  $\mathcal{A}_{\zeta_1}$ . Concerning the measurable hull, see [Roh52, §3.3] and [CK12, §1.4]. Moreover, compare [Hae73, Proposition 11].

The map from the trajectory space  $\Omega$  to the Poisson–Furstenberg boundary  $B$  that assigns to every trajectory  $x \in \Omega$  the respective element of the partition  $\zeta_1$  is called the boundary map  $\text{bnd} : \Omega \rightarrow B$ .

Note that the above is not the only possible definition of the Poisson–Furstenberg boundary, further equivalent ones are given in [KV83]. One important feature of the Poisson–Furstenberg boundary is that it can be used to describe all bounded harmonic functions on the state space  $X$ .

**Definition 4.2 (“harmonic function”)** *Assume we are given a countable state space  $X$  and transition probabilities  $p : X \times X \rightarrow [0, 1]$  as introduced in Section 3.1. A function  $\varphi : X \rightarrow \mathbb{R}$  is called harmonic if for every element  $x \in X$  the equation  $\varphi(x) = \sum_{y \in X} p(x, y)\varphi(y)$  holds. In other words, being at  $x \in X$ , the value of  $\varphi$  today is exactly as large as the expected value of  $\varphi$  tomorrow.*

The initial probability measure of a random walk is denoted by  $\vartheta : X \rightarrow [0, 1]$ . First, we pick some reference measure  $\vartheta$  with  $\text{supp}(\vartheta) = X$ . Then, we consider the random walk  $Z = (Z_0, Z_1, \dots)$  that starts according to  $\vartheta$ , has probability measure  $\mathbb{P}_\vartheta$  and Poisson–Furstenberg boundary  $(B, \mathcal{B}, \nu_\vartheta)$ .

All other initial probability measures, in particular the Dirac measures  $\delta_x$  at points  $x \in X$ , are absolutely continuous with respect to  $\vartheta$ . Therefore, the measures  $\mathbb{P}_x := \mathbb{P}_{\delta_x}$  are absolutely continuous with respect to  $\mathbb{P}_\vartheta$ , which implies that we may endow  $(B, \mathcal{B})$  with measures  $\nu_x := \nu_{\delta_x}$  in order to obtain the respective Poisson–Furstenberg boundaries.

From this point of view, it would have made sense to define the Poisson–Furstenberg boundary as a measurable space  $(B, \mathcal{B})$  endowed with a family of measures. A first step decomposition shows that for every two points  $x, y \in X$  the equation  $\nu_x = \sum_{y \in X} p(x, y) \cdot \nu_y$  holds. Hence, given an essentially bounded function  $f$  mapping from the Poisson–Furstenberg boundary  $(B, \mathcal{B}, \nu_\vartheta)$  to the real numbers  $\mathbb{R}$ , we can construct a bounded harmonic function  $\varphi : X \rightarrow \mathbb{R}$  given by the Poisson integral representation formula  $\varphi(x) := \int f d\nu_x$ .

There is also a way back from  $\varphi$  to  $f$  using martingale convergence so that, in the end, one obtains a one-to-one correspondence, even an isometry of Banach spaces, between the space  $L^\infty(B, \mathcal{B}, \nu_\vartheta)$  of equivalence classes of essentially bounded functions and the space  $H^\infty(X, \mu)$  of bounded harmonic functions, see e. g. [Kai96, Section 2.1].

### 4.3 Classical results about triviality and geometric identification

Given a random walk, be it on a generic state space or on a group, a challenging problem is to decide whether the Poisson–Furstenberg boundary is trivial or not. In the latter case, one may wonder how to

identify it geometrically. We shall only outline a few results about the Poisson–Furstenberg boundary of random walks on countable groups. A recent survey has been given by Erschler in [Ers10].

As before, let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a countable group  $G$  driven by the probability measure  $\mu$ . We assume that the support  $\text{supp}(\mu)$  generates  $G$  as a semigroup, see Section 3.1.

If  $G$  is abelian, then the Poisson–Furstenberg boundary is always trivial, see [Bla55] and [CD60]. The same holds true for all groups of polynomial growth, and for groups of subexponential growth endowed with a probability measure  $\mu$  with finite first moment. For the special case of probability measures with finite support, see [Ave74], and for the general case, see e. g. [KW02, Theorem 5.3] and [Ers04, §4]. Moreover, it was shown in [Ers04], that the assumption of finite first moment cannot be dropped.

If  $G$  is amenable, then one can show that there is at least one symmetric probability measure  $\mu$  such that the Poisson–Furstenberg boundary is trivial, see the conjecture in [Fur73, §9]. The proof of the conjecture has been announced in [VK79, Theorem 4] and given in [Ros81] and [KV83]. In case of the Baumslag–Solitar group  $G = \text{BS}(1, 2)$ , the Poisson–Furstenberg boundary may or may not be trivial depending on the vertical drift, see Definition 3.6. More precisely, for random walks on  $G = \text{BS}(1, 2)$  with finite first moment the Poisson–Furstenberg boundary is isomorphic to  $\mathbb{R}$  for  $\delta < 0$  and trivial for  $\delta = 0$  and isomorphic to  $\mathbb{Q}_2$  for  $\delta > 0$ , see [Kai91, Theorem 5.1]. We may think of  $\mathbb{Q}_2$  as the space of upper ends of the corresponding Bass–Serre tree  $\mathbb{T}$ . Further results about random walks on rational affinities are given in [Bro06]. For the Poisson–Furstenberg boundary of lamplighter random walks, see [VK79], [KV83], [LP15], and also [Sav10].

If  $G$  is non-amenable, then the Poisson–Furstenberg boundary is always non-trivial, see [Fur73, §9]. In particular, this implies that the Poisson–Furstenberg boundary of random walks on non-amenable Baumslag–Solitar groups can never be trivial, also when  $\delta = 0$ .

#### 4.4 Kaimanovich’s strip criterion

Kaimanovich’s strip criterion is a tool for identifying the Poisson–Furstenberg boundary geometrically. The strategy is to choose a suitable  $\mu$ -boundary as a candidate. Our one will be given in terms of the boundaries  $\partial\mathbb{H}$  and  $\partial\mathbb{T}$ . The strip criterion then enables us to prove that our candidate is indeed isomorphic to the Poisson–Furstenberg boundary. Let us first recall the strip criterion. For a proof we refer to [Kai00, §6.4].

**Theorem 4.3 (“strip criterion”)** *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a countable group  $G$  driven by a probability measure  $\mu$  with finite entropy  $H(\mu)$ . Moreover, let  $(B_-, \mathcal{B}_-, \nu_-)$  and  $(B_+, \mathcal{B}_+, \nu_+)$  be  $\check{\mu}$ - and  $\mu$ -boundaries, respectively. If there exist a gauge  $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \dots)$  on  $G$  with associated gauge function  $|\cdot| = |\cdot|_{\mathcal{G}}$  and a measurable  $G$ -equivariant map  $\mathcal{S}$  assigning to pairs of points  $(b_-, b_+) \in B_- \times B_+$  non-empty strips  $\mathcal{S}(b_-, b_+) \subseteq G$  such that for every  $g \in G$  and  $\nu_- \otimes \nu_+$ -almost every  $(b_-, b_+) \in B_- \times B_+$*

$$\frac{1}{n} \ln (\text{card}(\mathcal{S}(b_-, b_+)g \cap \mathcal{G}_{|Z_n|})) \xrightarrow{n \rightarrow \infty} 0 \text{ in probability,}$$

*then the  $\mu$ -boundary  $(B_+, \mathcal{B}_+, \nu_+)$  is maximal.*

**Remark 4.4** *The proof shows that it is not even necessary to verify the convergence for every  $g \in G$ . It suffices to consider the special case  $g = 1$  as long as we can ensure that a random strip contains the identity element  $1 \in G$  with positive probability, i. e. that  $\nu_- \otimes \nu_+ \{(b_-, b_+) \in B_- \times B_+ \mid 1 \in \mathcal{S}(b_-, b_+)\} > 0$ .*

The following four notions have not yet been introduced.

**(a) “entropy”** — The entropy of the probability measure  $\mu$  is the expected amount of information contained in the outcome of a random variable that is distributed according to  $\mu$ . More precisely, it is the real number given by  $H(\mu) := \sum_{g \in G} -\log_2(\mu(g)) \cdot \mu(g)$ . Here, as usual, one defines  $-\log_2(0) \cdot 0 := 0$ . For us, the assumption of finite entropy will be no issue because Baumslag–Solitar groups are finitely generated and the increments under investigation have finite first moment. This implies that their probability measures  $\mu$  have finite entropy, as shown by the following well-known lemma.

**Lemma 4.5** *Let  $G$  be a finitely generated group<sup>4</sup> and let  $\mu : G \rightarrow [0, 1]$  be a probability measure. If a random variable  $X : \Omega \rightarrow G$  distributed according to  $\mu$  has finite first moment, then  $\mu$  has finite entropy.*

*Proof.* Let  $S \subseteq G$  be a non-empty finite generating set. Moreover, let  $b := 2 \cdot |S| + 1$ , whence  $b \geq 3$ . In this proof, we shall use the shorthand notation  $d$  instead of  $d_S$  to denote the word metric on  $G$ . We have to show that the entropy  $H(\mu) = \sum_{g \in G} -\log_2(\mu(g)) \cdot \mu(g)$  is finite. First, we change the base of the logarithm

$$H(\mu) = \sum_{g \in G} -\log_2(\mu(g)) \cdot \mu(g) = \log_2(b) \cdot \sum_{g \in G} -\log_b(\mu(g)) \cdot \mu(g),$$

and split the summands appropriately

$$\dots = \log_2(b) \cdot \left( -\log_b(\mu(1)) \cdot \mu(1) + \sum_{\substack{g \in G \setminus \{1\} \\ \mu(g) \leq b^{-d(1,g)}}} -\log_b(\mu(g)) \cdot \mu(g) + \sum_{\substack{g \in G \setminus \{1\} \\ \mu(g) > b^{-d(1,g)}}} -\log_b(\mu(g)) \cdot \mu(g) \right).$$

Then, we recall that the function  $x \mapsto -\log_b(x) \cdot x$  is increasing on the interval  $[0, \frac{1}{e}]$ , and conclude that

$$\begin{aligned} \sum_{\substack{g \in G \setminus \{1\} \\ \mu(g) \leq b^{-d(1,g)}}} -\log_b(\mu(g)) \cdot \mu(g) &\leq \sum_{\substack{g \in G \setminus \{1\} \\ \mu(g) \leq b^{-d(1,g)}}} -\log_b(b^{-d(1,g)}) \cdot b^{-d(1,g)} = \sum_{\substack{g \in G \setminus \{1\} \\ \mu(g) \leq b^{-d(1,g)}}} d(1, g) \cdot b^{-d(1,g)} \\ &\leq \sum_{g \in G} d(1, g) \cdot b^{-d(1,g)} \leq \sum_{n=0}^{\infty} (2 \cdot |S|)^n \cdot n \cdot b^{-n} < \infty. \end{aligned}$$

On the other hand, since  $X$  has finite first moment,

$$\sum_{\substack{g \in G \setminus \{1\} \\ \mu(g) > b^{-d(1,g)}}} -\log_b(\mu(g)) \cdot \mu(g) \leq \sum_{\substack{g \in G \setminus \{1\} \\ \mu(g) > b^{-d(1,g)}}} d(1, g) \cdot \mu(g) \leq \sum_{g \in G} d(1, g) \cdot \mu(g) < \infty.$$

So, both sums are finite, whence  $H(\mu)$  must be finite, too.  $\square$

**(b) “ $\mu$ -boundary”** — Two equivalent definitions of a  $\mu$ -boundary can be found in [Kai00, §1.5]. For us, it suffices to record that every Lebesgue–Rohlin space  $(B_+, \mathcal{B}_+, \nu_+)$  endowed with a left  $G$ -action and a boundary map  $\text{bnd}_+ : \Omega \rightarrow B_+$  that is ① measurable,<sup>5</sup> ②  $\sim$ -invariant, and ③  $G$ -equivariant is a  $\mu$ -boundary.

<sup>4</sup>This assumption is necessary. For example, imagine the group  $(\mathbb{Q}, +)$  and a probability measure  $\mu$  with infinite entropy supported on the generating set  $S := \{1, -1, \frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \dots\}$ . Then,  $X$  has finite first moment with respect to the word metric  $d_S : \mathbb{Q} \times \mathbb{Q} \rightarrow \mathbb{R}$ , but  $\mu$  has infinite entropy.

<sup>5</sup>Here, the term measurable means being a measurable homomorphism (=measurable and measure preserving map) between Lebesgue–Rohlin spaces. There is a natural correspondence between measurable homomorphisms and measurable partitions of their domain, see [Roh52, §3.2] and [Hae73, p. 255, Remark] for details.



(c) “ **$\check{\mu}$ -boundary**” — While  $\mu$  is the probability measure driving the random walk, the symbol  $\check{\mu}$  denotes the reflected probability measure given by  $\check{\mu}(g) := \mu(g^{-1})$ . Accordingly, a  $\check{\mu}$ -boundary is a space that satisfies the requirements of a  $\mu$ -boundary when replacing  $\mu$  by  $\check{\mu}$ .

(d) “**gauge**” — A gauge  $\mathcal{G}$  is an exhaustion  $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \dots)$  of the group  $G$ , i. e. a sequence of subsets  $\mathcal{G}_k \subseteq G$  which is increasing  $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \dots$  and whose union  $\mathcal{G}_1 \cup \mathcal{G}_2 \cup \dots$  is the whole group  $G$ . Given a gauge  $\mathcal{G}$  and an element  $g \in G$ , we may ask for the minimal index  $k \in \mathbb{N}$  with the property that  $g \in \mathcal{G}_k$ . This index is the value of the associated gauge function  $|\cdot| = |\cdot|_{\mathcal{G}}$  at  $g$ .

**Remark 4.6** *Kaimanovich distinguishes between various kinds of gauges, see [Kai00]. For example, a gauge  $\mathcal{G}$  is subadditive if any two group elements  $g_1, g_2 \in G$  satisfy  $|g_1 g_2| \leq |g_1| + |g_2|$  and it is temperate if all gauge sets  $\mathcal{G}_k$  are finite and grow at most exponentially. Even though these two properties do play a crucial role in the corollaries to the strip criterion given in [Kai00, §6.5], they are not required in the strip criterion itself. And, in fact, not all of our gauges will have these two properties.*

(e) “**measurable strips**” — The power set  $\{0, 1\}^G$  is naturally endowed with the product  $\sigma$ -algebra, which enables us to talk about measurability of the map  $\mathcal{S} : B_- \times B_+ \rightarrow \{0, 1\}^G$ .

(f) “**maximal**” — The Poisson–Furstenberg boundary  $(B, \mathcal{B}, \nu)$  inherits a left  $G$ -action from the space of trajectories, and every  $\mu$ -boundary  $(B_+, \mathcal{B}_+, \nu_+)$  is a  $G$ -invariant measurable quotient of  $(B, \mathcal{B}, \nu)$ .

$$\begin{array}{ccc}
 & (\Omega, \mathcal{A}, \mathbb{P}) & \\
 \text{bnd} \swarrow & & \searrow \text{bnd}_+ \\
 (B, \mathcal{B}, \nu) & \xrightarrow{\pi} & (B_+, \mathcal{B}_+, \nu_+)
 \end{array}$$

A  $\mu$ -boundary  $(B_+, \mathcal{B}_+, \nu_+)$  is called maximal if the projection  $\pi : (B, \mathcal{B}, \nu) \rightarrow (B_+, \mathcal{B}_+, \nu_+)$  is a measurable isomorphism between Lebesgue–Rohlin spaces.

## 5 Identification of the Poisson–Furstenberg boundary

We still assume that  $1 \leq p < q$  and consider a random walk  $Z = (Z_0, Z_1, \dots)$  on  $G = \text{BS}(p, q)$ . Moreover, recall the abbreviations  $A_g := \text{Im}(\pi_{\mathbb{H}}(g))$  and  $B_g := \text{Re}(\pi_{\mathbb{H}}(g))$  introduced in Section 3.3.

### 5.1 Convergence to the boundary of the hyperbolic plane

The following lemmas concern the behaviour of the projections  $\pi_{\mathbb{H}}(Z_n)$ . They seem to be well-known and we do not claim originality. But, for the sake of completeness, we give rigorous proofs.

**Lemma 5.1** *Assume that  $\ln(A_{X_1})$  has finite first moment. If the vertical drift is positive, i. e.  $\delta > 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a. s. to  $\infty \in \partial\mathbb{H}$ .*

*Proof.* We can use the strong law of large numbers to obtain

$$\frac{\lambda(Z_n)}{n} = \frac{\lambda(X_1) + \dots + \lambda(X_n)}{n} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} \mathbb{E}(\lambda(X_1)) = \delta > 0.$$



Hence, the projections  $\lambda(Z_n)$  tend a. s. to infinity and, by Remark 3.5, the imaginary parts  $A_{Z_n}$  do. This, of course, implies that the absolute values  $|\pi_{\mathbb{H}}(Z_n)|$  tend a. s. to infinity. Now, we can use Lemma 2.8 to complete the proof.  $\square$

**Lemma 5.2** *Assume that both  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$  have finite first moment. If the vertical drift is negative, i. e.  $\delta < 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a. s. to a random element  $r \in \partial\mathbb{H} \setminus \{\infty\}$ .*

*Proof.* Note that the argument given in the proof of Lemma 5.1 can be adapted to show that the imaginary parts  $A_{Z_n}$  converge a. s. to 0, whence we only need to understand the behaviour of the real parts  $B_{Z_n}$ . The equation  $\pi_{\mathbb{H}}(Z_n) = A_{Z_n} \cdot i + B_{Z_n}$  yields  $\pi_{\text{Aff}^+(\mathbb{R})}(Z_n)(z) = A_{Z_n} \cdot z + B_{Z_n}$ , and in light of the multiplication in  $\text{Aff}^+(\mathbb{R})$  we obtain

$$\begin{aligned} \pi_{\mathbb{H}}(Z_n) &= \pi_{\text{Aff}^+(\mathbb{R})}(Z_n)(i) = \pi_{\text{Aff}^+(\mathbb{R})}(X_1 \cdots X_n)(i) \\ &= (\pi_{\text{Aff}^+(\mathbb{R})}(X_1) \circ \cdots \circ \pi_{\text{Aff}^+(\mathbb{R})}(X_n))(i) \\ &= A_{X_1} \cdots A_{X_n} \cdot i + \sum_{k=1}^n A_{X_1} \cdots A_{X_{k-1}} \cdot B_{X_k}. \end{aligned}$$

Hence, the real parts  $B_{Z_n}$  are partial sums of the infinite series  $\sum_{k=1}^{\infty} C_k$  with  $C_k := A_{X_1} \cdots A_{X_{k-1}} \cdot B_{X_k}$ . In order to verify a. s. convergence of the latter, we apply Cauchy's root test,

$$|C_k|^{\frac{1}{k}} \leq \exp \left( \ln \left( \frac{q}{p} \right) \cdot \underbrace{\frac{\lambda(X_1) + \cdots + \lambda(X_{k-1})}{k-1}}_{\rightarrow \mathbb{E}(\lambda(X_1)) = \delta < 0 \text{ a. s.}} \cdot \underbrace{\frac{k-1}{k}}_{\rightarrow 1} \right) \cdot \exp \left( \underbrace{\frac{\ln(1 + |B_{X_k}|)}{k}}_{\rightarrow 0 \text{ a. s.}} \right) \xrightarrow[k \rightarrow \infty]{\text{a. s.}} \left( \frac{q}{p} \right)^{\delta} < 1.$$

For the convergence claimed in the first factor we can use the strong law of large numbers, for the one claimed in the second factor the Borel–Cantelli Lemma. Indeed, let us write  $Q_k$  for the quotient  $\frac{1}{k} \cdot \ln(1 + |B_{X_k}|)$ . In order to show that  $Q_k \rightarrow 0$  a. s., recall that  $\ln(1 + |B_{X_1}|)$  has finite first moment. So, for every  $\varepsilon > 0$  we may estimate

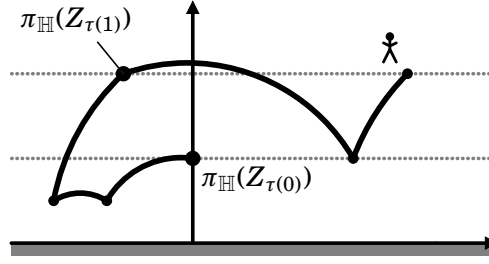
$$\sum_{k=1}^{\infty} \mathbb{P}(\{\omega \in \Omega \mid Q_k(\omega) > \varepsilon\}) \leq \sum_{k=1}^{\infty} \mathbb{P} \left( \left\{ \omega \in \Omega \mid \left\lceil \frac{\ln(1 + |B_{X_1}(\omega)|)}{\varepsilon} \right\rceil \geq k \right\} \right) = \mathbb{E} \left( \left\lceil \frac{\ln(1 + |B_{X_1}|)}{\varepsilon} \right\rceil \right).$$

Therefore, the Borel–Cantelli Lemma yields  $\mathbb{P}(\{\omega \in \Omega \mid \exists \text{ infinitely many } k \in \mathbb{N} \text{ such that } Q_k(\omega) > \varepsilon\}) = 0$ . Replacing  $\varepsilon$  by  $1, \frac{1}{2}, \frac{1}{3}, \dots$ , we obtain a countable family of null sets whose union is, of course, again a null set that consists of all  $\omega \in \Omega$  with  $Q_k(\omega) \not\rightarrow 0$ . Hence,  $Q_k \rightarrow 0$  a. s., see also [Kle14, Exercise 5.1.3]. So, we have finally convinced ourselves that  $\limsup_{k \rightarrow \infty} |C_k|^{\frac{1}{k}} < 1$  a. s., whence  $\sum_{k=1}^{\infty} C_k$  converges a. s. to a random element  $r \in \mathbb{R}$ .  $\square$

The natural question that remains is the one asking for the driftless case. An answer has been given by Brofferio in [Bro03, Theorem 1]. It says that under the same mild assumptions, namely that  $\ln(A_{X_1})$  and  $\ln(1 + |B_{X_1}|)$  have finite first moment, the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a. s. to  $\infty \in \partial\mathbb{H}$ . But, for us, a result of slightly different flavour will be of relevance.

**Lemma 5.3** *Assume that  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment. If there is no vertical drift, i. e.  $\delta = 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  have sublinear speed, i. e.*

$$\frac{d_{\mathbb{H}}(\pi_{\mathbb{H}}(Z_0), \pi_{\mathbb{H}}(Z_n))}{n} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} 0.$$


 Figure 6: The first ladder times  $\tau(0)$  and  $\tau(1)$ .

The proof is based on ideas that go back to Élie in [Éli82, Lemme 5.49] and have also been used by Cartwright, Kaimanovich, and Woess in [CKW94, Proposition 4b]. We first adapt these ideas to our situation by stating and proving Lemma 5.4, and then deduce Lemma 5.3.

By assumption, there is no vertical drift so that the pointwise projection  $\lambda(Z) = (\lambda(Z_0), \lambda(Z_1), \dots)$  is recurrent, see Pólya's Theorem for the simple random walk and the Chung–Fuchs Theorem in [CF51] for the general case. In particular, we know that there exists a.s. a strictly increasing sequence  $\tau(0), \tau(1), \dots$  given by  $\tau(0) := 0$  and by  $\tau(n) := \inf\{k \in \mathbb{N} \mid \tau(n-1) < k \text{ and } \lambda(Z_{\tau(n-1)}) < \lambda(Z_k)\}$  for all  $n \in \mathbb{N}$ . We call  $\tau(n)$  the  $n$ -th ladder time, see Figure 6 for an illustration of the first ladder times  $\tau(0)$  and  $\tau(1)$ . The following lemma concerns the random variable  $\ln(1 + \sum_{k=1}^{\tau} |B_{X_k}|)$  with  $\tau := \tau(1)$ .

**Lemma 5.4** *Under the same assumptions as in Lemma 5.3, namely that  $\ln(A_{X_1})$  has finite second moment, there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment, and there is no vertical drift, i. e.  $\delta = 0$ , the random variable  $\ln(1 + \sum_{k=1}^{\tau} |B_{X_k}|)$  has finite first moment.*

*Proof.* Adapting the proof of [Éli82, Lemme 5.49], we begin with some preliminaries. Pick an  $\varepsilon > 0$  that satisfies the requirements of Lemma 5.4 and let  $\beta := \frac{1}{2+\varepsilon}$ . Since  $\ln(A_{X_1})$  has finite second moment, we know that also  $\lambda(X_1)$  has finite second moment and  $\mathbb{P}(\{\omega \in \Omega \mid \tau(\omega) > k\}) \sim \text{const} \cdot k^{-\frac{1}{2}}$ , see [Éli82, §5.44] referring to [Fel71, p. 415]. Using this asymptotics, we obtain

$$\int \tau^\beta d\mathbb{P} \leq \int \lceil \tau^\beta \rceil d\mathbb{P} = \sum_{k=1}^{\infty} \mathbb{P}\left(\left\{\omega \in \Omega \mid \lceil \tau(\omega)^\beta \rceil \geq k\right\}\right) = \sum_{k=0}^{\infty} \underbrace{\mathbb{P}\left(\left\{\omega \in \Omega \mid \tau(\omega) > k^{\frac{1}{\beta}}\right\}\right)}_{\sim \text{const} \cdot k^{-(1+\frac{\varepsilon}{2})}}.$$

In particular, there is a  $k_0 \in \mathbb{N}$  such that for all  $k \geq k_0$  the summands  $\mathbb{P}(\{\omega \in \Omega \mid \tau(\omega) > k^{\frac{1}{\beta}}\})$  are strictly smaller than  $k^{-(1+\frac{\varepsilon}{4})}$ . Since  $\sum_{k=k_0}^{\infty} k^{-(1+\frac{\varepsilon}{4})} < \infty$ , we know that  $\int \tau^\beta d\mathbb{P} < \infty$ . Moreover, note that, by construction of the ladder times  $\tau(0), \tau(1), \dots$ , the differences  $\tau(1) - \tau(0), \tau(2) - \tau(1), \dots$  are i. i. d., whence the fact that  $0 < \beta < 1$ , which implies that  $(x + y)^\beta \leq x^\beta + y^\beta$ , and the strong law of large numbers yield

$$\begin{aligned} \frac{\tau(n)^\beta}{n} &\leq \frac{(\tau(1) - \tau(0))^\beta + \dots + (\tau(n) - \tau(n-1))^\beta}{n} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} \mathbb{E}(\tau^\beta), \\ &\Rightarrow \limsup_{n \rightarrow \infty} \frac{\tau(n)^\beta}{n} < \infty \text{ a. s.} \end{aligned} \quad (*)$$

Now, we are prepared for the main argument. Recall that we aim to show that  $\ln(1 + \sum_{k=1}^{\tau} |B_{X_k}|)$  has finite first moment. The sums  $\sum_{k=\tau(0)+1}^{\tau(1)} |B_{X_k}|$ ,  $\sum_{k=\tau(1)+1}^{\tau(2)} |B_{X_k}|$ ,  $\dots$  are i. i. d. and non-negative with the

$$\int \ln \left( 1 + \sum_{k=1}^{\tau} |B_{X_k}| \right) d\mathbb{P} < \infty \iff \underbrace{\limsup_{n \rightarrow \infty} \left( \sum_{k=\tau(n-1)+1}^{\tau(n)} |B_{X_k}| \right)^{\frac{1}{n}}}_{=: K} < \infty \text{ a. s.}$$

$$K \leq \limsup_{n \rightarrow \infty} \exp \left( \frac{\ln \left( 1 + \sum_{k=1}^{\tau(n)} |B_{X_k}| \right)}{n} \right) \leq \exp \left( \underbrace{\limsup_{n \rightarrow \infty} \frac{\ln \left( 1 + \sum_{k=1}^{\tau(n)} |B_{X_k}| \right)}{\tau(n)^\beta}_{=: L} \cdot \underbrace{\limsup_{n \rightarrow \infty} \frac{\tau(n)^\beta}{n}}_{< \infty \text{ a.s. } (*)} \right).$$
$$\begin{aligned} L &= \limsup_{n \rightarrow \infty} \frac{\ln \left( 1 + \sum_{k=1}^{\tau(n)} |B_{X_k}| \right)}{\tau(n)^\beta} \leq \limsup_{n \rightarrow \infty} \frac{\ln \left( 1 + \tau(n) \cdot \max_{1 \leq k \leq \tau(n)} \{ |B_{X_k}| \} \right)}{\tau(n)^\beta} \\ &\leq \underbrace{\limsup_{n \rightarrow \infty} \frac{\ln(\tau(n))}{\tau(n)^\beta}}_{=0} + \limsup_{n \rightarrow \infty} \frac{\ln \left( 1 + \max_{1 \leq k \leq \tau(n)} \{ |B_{X_k}| \} \right)}{\tau(n)^\beta} \\ &= \limsup_{n \rightarrow \infty} \left( \frac{\max_{1 \leq k \leq \tau(n)} \left\{ \ln \left( 1 + |B_{X_k}| \right)^{\frac{1}{\beta}} \right\}}{\tau(n)} \right)^\beta \leq \limsup_{n \rightarrow \infty} \left( \underbrace{\frac{\sum_{k=1}^{\tau(n)} \ln \left( 1 + |B_{X_k}| \right)^{\frac{1}{\beta}}}{\tau(n)}}_{=: M_n} \right)^\beta. \end{aligned}$$
$$M_n \xrightarrow[n \rightarrow \infty]{a.s.} \mathbb{E} \left( \ln(1 + |B_{X_1}|)^{\frac{1}{\beta}} \right).$$

*Proof of Lemma 5.3.* Recall from [Éli82, §5.44] and [Fel71, p. 415] that  $\mathbb{P}(\{\omega \in \Omega \mid \tau(\omega) > k\}) \sim \text{const} \cdot k^{-\frac{1}{2}}$  with a strictly positive constant. In particular, there is a  $k_0 \in \mathbb{N}$  such that for all  $k \geq k_0$  the summands  $\mathbb{P}(\{\omega \in \Omega \mid \tau(\omega) > k\})$  are strictly larger than  $k^{-1}$  and we obtain

$$\int \tau \, d\mathbb{P} = \sum_{k=1}^{\infty} \mathbb{P}(\{\omega \in \Omega \mid \tau(\omega) \geq k\}) = \sum_{k=0}^{\infty} \mathbb{P}(\{\omega \in \Omega \mid \tau(\omega) > k\}) \geq \sum_{k=k_0}^{\infty} k^{-1} = \infty.$$

$$\frac{\tau(n)}{n} = \frac{(\tau(1) - \tau(0)) + (\tau(2) - \tau(1)) + \dots + (\tau(n) - \tau(n-1))}{n} \xrightarrow[\text{a. s.}]{n \rightarrow \infty} \infty \quad \text{and} \quad \frac{n}{\tau(n)} \xrightarrow[\text{a. s.}]{n \rightarrow \infty} 0.$$

<sup>6</sup>This can be done by truncating the random variables, see e. g. [Rou14, p. 309, Lemma 6].

This can be used to estimate the distance between  $\pi_{\mathbb{H}}(Z_0)$  and  $\pi_{\mathbb{H}}(Z_n)$  from above. First, for every  $n \in \mathbb{N}_0$  let  $m = m(n) \in \mathbb{N}_0$  be the unique element with  $\tau(m) \leq n < \tau(m+1)$ . This element exists a. s. because the ladder times  $0 = \tau(0) < \tau(1) < \dots$  do. Now, we may estimate

$$\frac{d_{\mathbb{H}}(\pi_{\mathbb{H}}(Z_0), \pi_{\mathbb{H}}(Z_n))}{n} \leq \underbrace{\frac{d_{\mathbb{H}}(i, A_{Z_{\tau(m)}} \cdot i)}{n}}_{\textcircled{1}} + \underbrace{\frac{d_{\mathbb{H}}(A_{Z_{\tau(m)}} \cdot i, A_{Z_{\tau(m)}} \cdot i + B_{Z_n})}{n}}_{\textcircled{2}} + \underbrace{\frac{d_{\mathbb{H}}(A_{Z_{\tau(m)}} \cdot i + B_{Z_n}, A_{Z_n} \cdot i + B_{Z_n})}{n}}_{\textcircled{3}}.$$

The numbers refer to Figure 7. We will consider the three summands separately and show that each of them converges a. s. to 0. For  $\textcircled{1}$  and  $\textcircled{3}$  this is straightforward. Indeed,

$$\textcircled{1} = \frac{|\ln(A_{Z_{\tau(m)}})|}{n} \leq \frac{|\ln(A_{Z_{\tau(m)}})|}{\tau(m)} = \ln\left(\frac{q}{p}\right) \cdot \left| \frac{\lambda(X_1) + \dots + \lambda(X_{\tau(m)})}{\tau(m)} \right| \xrightarrow[\text{a. s.}]{n \rightarrow \infty} \ln\left(\frac{q}{p}\right) \cdot |\mathbb{E}(\lambda(X_1))| = \ln\left(\frac{q}{p}\right) \cdot |\delta| = 0$$

and similarly

$$\textcircled{3} \leq \frac{d_{\mathbb{H}}(A_{Z_{\tau(m)}} \cdot i, i)}{n} + \frac{d_{\mathbb{H}}(i, A_{Z_n} \cdot i)}{n} = \textcircled{1} + \frac{|\ln(A_{Z_n})|}{n} = \textcircled{1} + \ln\left(\frac{q}{p}\right) \cdot \left| \frac{\lambda(X_1) + \dots + \lambda(X_n)}{n} \right| \xrightarrow[\text{a. s.}]{n \rightarrow \infty} 0.$$

For  $\textcircled{2}$  recall from the proof of Lemma 5.2 that  $B_{Z_n} = \sum_{k=1}^n A_{X_1} \cdot \dots \cdot A_{X_{k-1}} \cdot B_{X_k}$  and observe that for every  $m, \ell \in \mathbb{N}_0$  with  $\tau(m) \leq \ell \leq \tau(m+1)$  the following estimate holds

$$\begin{aligned} \frac{|B_{Z_\ell} - B_{Z_{\tau(m)}}|}{A_{Z_{\tau(m)}}} &\leq \frac{A_{X_1} \cdot \dots \cdot A_{X_{\tau(m)}} \cdot \sum_{k=\tau(m)+1}^{\ell} A_{X_{\tau(m)+1}} \cdot \dots \cdot A_{X_{k-1}} \cdot |B_{X_k}|}{A_{X_1} \cdot \dots \cdot A_{X_{\tau(m)}}} \\ &= \sum_{k=\tau(m)+1}^{\ell} \underbrace{A_{X_{\tau(m)+1}} \cdot \dots \cdot A_{X_{k-1}}}_{\leq 1} \cdot |B_{X_k}| \leq \sum_{k=\tau(m)+1}^{\ell} |B_{X_k}| \quad \text{a. s.} \end{aligned} \quad (*)$$

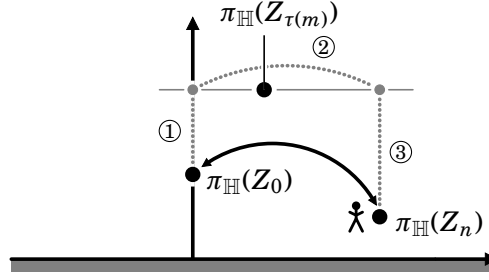
Hence, using that  $A_{Z_{\tau(0)}} < A_{Z_{\tau(1)}} < \dots < A_{Z_{\tau(m)}}$  and  $n < \tau(m+1)$ , we obtain

$$\begin{aligned} \textcircled{2} &= \frac{1}{n} \cdot \operatorname{arcosh} \left( 1 + \frac{1}{2} \cdot \left( \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right)^2 \right) = \frac{1}{n} \cdot \ln \left( 1 + \frac{1}{2} \cdot \left( \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right)^2 + \sqrt{\left( 1 + \frac{1}{2} \cdot \left( \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right)^2 \right)^2 - 1} \right) \\ &\leq \frac{1}{n} \cdot \left( \ln(2) + \ln \left( 1 + \left( \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right)^2 \right) \right) \leq \frac{1}{n} \cdot \left( \ln(2) + 2 \cdot \ln \left( 1 + \frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}} \right) \right) \\ &\leq \frac{1}{n} \cdot \left( \ln(2) + 2 \cdot \ln \left( 1 + \frac{|B_{Z_{\tau(1)}} - B_{Z_{\tau(0)}}|}{A_{Z_{\tau(0)}}} + \frac{|B_{Z_{\tau(2)}} - B_{Z_{\tau(1)}}|}{A_{Z_{\tau(1)}}} + \dots + \frac{|B_{Z_{\tau(m)}} - B_{Z_{\tau(m-1)}}|}{A_{Z_{\tau(m-1)}}} + \frac{|B_{Z_n} - B_{Z_{\tau(m)}}|}{A_{Z_{\tau(m)}}} \right) \right), \end{aligned}$$

which allows us to apply  $(*)$  and finally conclude

$$\begin{aligned} \dots &\leq \frac{1}{n} \cdot \left( \ln(2) + 2 \cdot \ln \left( 1 + \sum_{k=1}^n |B_{X_k}| \right) \right) \leq \frac{1}{n} \cdot \left( \ln(2) + 2 \cdot \ln \left( 1 + \sum_{k=1}^{\tau(m+1)} |B_{X_k}| \right) \right) \\ &\leq \underbrace{\frac{\ln(2)}{n}}_{\rightarrow 0} + 2 \cdot \underbrace{\frac{\ln \left( 1 + \sum_{k=\tau(0)+1}^{\tau(1)} |B_{X_k}| \right) + \dots + \ln \left( 1 + \sum_{k=\tau(m)+1}^{\tau(m+1)} |B_{X_k}| \right)}{m+1}}_{\rightarrow \mathbb{E}(\ln(1 + \sum_{k=1}^{\tau} |B_{X_k}|)) \text{ a. s. by Lemma 5.4}} \cdot \underbrace{\frac{m+1}{\tau(m)}}_{\rightarrow 0 \text{ a. s.}} \cdot \underbrace{\frac{\tau(m)}{n}}_{\leq 1 \text{ a. s.}} \xrightarrow[\text{a. s.}]{n \rightarrow \infty} 0. \end{aligned}$$

□


 Figure 7: Estimate of the distance between  $\pi_{\mathbb{H}}(Z_0)$  and  $\pi_{\mathbb{H}}(Z_n)$ .

## 5.2 Convergence to the space of ends of the Bass–Serre tree

Unlike the ones considered in Section 5.1, the projections  $\pi_{\mathbb{T}}(Z_n)$  do not need to satisfy the Markov property. Consider, for example, the random walk  $Z = (Z_0, Z_1, \dots)$  driven by the uniform measure on the standard generators and their formal inverses. Then, given  $\pi_{\mathbb{T}}(Z_{k-2}) = B$  and  $\pi_{\mathbb{T}}(Z_{k-1}) = a^{-1}B$ , the projection  $\pi_{\mathbb{T}}(Z_k)$  comes back to  $B$  with probability  $\frac{1}{4}$ . On the other hand, coming back to  $B$  in a single step would not be possible if the history was  $\pi_{\mathbb{T}}(Z_{k-3}) = B$  and  $\pi_{\mathbb{T}}(Z_{k-2}) = \pi_{\mathbb{T}}(Z_{k-1}) = a^{-1}B$ . Despite of this subtlety, the following lemmas yield almost sure convergence of the projections  $\pi_{\mathbb{T}}(Z_n)$  to a random end.

**Lemma 5.5** *Assume that  $X_1$  has finite first moment. If the vertical drift is different from 0, i. e.  $\delta \neq 0$ , then the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random end  $\xi \in \partial\mathbb{T}$ .*

By Lemma 3.4, the assumption that  $X_1$  has finite first moment implies that  $\ln(A_{X_1})$  has also finite first moment. For the proof of the Lemma 5.5, we give an argument using the notion of regular sequences, see [CKW94, §2.C]. One difference to [CKW94] is that we do not fix any particular end  $\omega \in \partial\mathbb{T}$ . Therefore, we replace the Busemann function  $h$ , which depends on the choice of  $\omega \in \partial\mathbb{T}$ , by the graph distance to the basepoint  $B$ . The other difference is that we work with the limit inferior instead of the limit in order to be prepared to deal with the driftless case, too.

*Proof of Lemma 5.5.* Let  $d_{\mathbb{T}}$  be the graph distance in the tree  $\mathbb{T}$ . Accordingly, the symbol  $|x|_{\mathbb{T}}$  denotes the graph distance  $d_{\mathbb{T}}(B, x)$  from the basepoint  $B$  to the vertex  $x$ . We call a sequence  $(x_0, x_1, \dots)$  of vertices regular if

$$\textcircled{1} \quad \liminf_{n \rightarrow \infty} \frac{|x_n|_{\mathbb{T}}}{n} > 0 \quad \text{and} \quad \textcircled{2} \quad \frac{d_{\mathbb{T}}(x_n, x_{n+1})}{n} \xrightarrow{n \rightarrow \infty} 0.$$

In order to prove Lemma 5.5, we pursue a two-step strategy. First, we show that every regular sequence converges to an end and, second, that the projections  $\pi_{\mathbb{T}}(Z_n)$  constitute a. s. a regular sequence.

Concerning the first part, we pick an arbitrary regular sequence  $(x_0, x_1, \dots)$  and claim that there is an end  $\xi \in \partial\mathbb{T}$  such that for every  $\varepsilon > 0$  the open ball  $B_{\varepsilon}(\xi) := \{x \in \widehat{\mathbb{T}} \mid d_{\widehat{\mathbb{T}}}(\xi, x) < \varepsilon\}$  contains infinitely many  $x_k$ . Assume there was no such end. Then, we know that for every  $\xi \in \partial\mathbb{T}$  there is an  $\varepsilon_1 = \varepsilon_1(\xi) > 0$  such that  $B_{\varepsilon_1}(\xi)$  contains only finitely many  $x_k$ , whence there is also an  $\varepsilon_2 = \varepsilon_2(\xi) > 0$  such that  $B_{\varepsilon_2}(\xi)$  does not contain any  $x_k$  at all. The open balls  $B_{\varepsilon_2}(\xi)$  with  $\xi \in \partial\mathbb{T}$  and the singletons  $\{x\}$  of vertices  $x$  form an open covering of  $\widehat{\mathbb{T}}$ . By compactness, it contains a finite subcovering. But since the constants  $\varepsilon_2 = \varepsilon_2(\xi) > 0$  have been chosen in such a way that the sequence  $(x_0, x_1, \dots)$  does not enter any of the open balls  $B_{\varepsilon_2}(\xi)$ , it must remain in a finite subset of the tree, which contradicts  $\textcircled{1}$ .

Next, we pick such an end  $\xi \in \partial\mathbb{T}$  and claim that the sequence  $(x_0, x_1, \dots)$  converges to  $\xi$ . Let  $\varepsilon > 0$ . We define  $\alpha := \frac{1}{3} \cdot \liminf_{n \rightarrow \infty} \frac{|x_n|_{\mathbb{T}}}{n}$ . By ①, the constant  $\alpha$  is strictly positive. Moreover, all but finitely many  $|x_n|_{\mathbb{T}}$  are strictly greater than  $2\alpha n$ . The elements in  $B_\varepsilon(\xi)$  are characterised by the property that the paths starting in  $B$  and representing them must have a certain finite initial piece. Let  $m$  be the length of this piece, i. e.  $m := \max\{0, \lfloor 1 - \log_2(\varepsilon) \rfloor\}$ , see Figure 8. By the above and by ②, there is an  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$  the inequalities  $|x_n|_{\mathbb{T}} > \alpha n + m$  and  $d_{\mathbb{T}}(x_n, x_{n+1}) < \alpha n$  hold. Since  $B_\varepsilon(\xi)$  contains infinitely many  $x_k$ , we can even find an  $n_1 \geq n_0$  such that  $x_{n_1} \in B_\varepsilon(\xi)$ . It turns out that not just for  $n_1$  but for all  $n \geq n_1$  we have  $x_n \in B_\varepsilon(\xi)$ . Indeed, if there was an  $n \geq n_1$  such that  $x_n \in B_\varepsilon(\xi)$  and  $x_{n+1} \notin B_\varepsilon(\xi)$ , we would know that

$$d_{\mathbb{T}}(x_n, x_{n+1}) \geq d_{\mathbb{T}}(x_n, x_n \wedge x_{n+1}) = |x_n|_{\mathbb{T}} - |x_n \wedge x_{n+1}|_{\mathbb{T}} > |x_n|_{\mathbb{T}} - m > \alpha n.$$

The latter, of course, contradicts  $d_{\mathbb{T}}(x_n, x_{n+1}) < \alpha n$ , see Figure 8. So, the two claims show that every regular sequence converges to an end. Concerning the second part, we aim to prove that

$$\textcircled{1} \quad \liminf_{n \rightarrow \infty} \frac{|\pi_{\mathbb{T}}(Z_n)|_{\mathbb{T}}}{n} > 0 \text{ a. s.} \quad \text{and} \quad \textcircled{2} \quad \frac{d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_n), \pi_{\mathbb{T}}(Z_{n+1}))}{n} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} 0.$$

Recall from Remark 3.5 that not just  $X_1$  and  $\ln(A_{X_1})$  have finite first moment but also  $\lambda(X_1)$  has. So, the strong law of large numbers yields

$$\frac{|\pi_{\mathbb{T}}(Z_n)|_{\mathbb{T}}}{n} \geq \frac{|\lambda(Z_n)|}{n} = \frac{|\lambda(X_1) + \dots + \lambda(X_n)|}{n} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} |\mathbb{E}(\lambda(X_1))| = |\delta| > 0$$

which implies that

$$\liminf_{n \rightarrow \infty} \frac{|\pi_{\mathbb{T}}(Z_n)|_{\mathbb{T}}}{n} > 0 \text{ a. s.}$$

Next, let  $S := \{a, b\} \subseteq G$  be the standard generating set. The numerators  $d(\pi_{\mathbb{T}}(Z_n), \pi_{\mathbb{T}}(Z_{n+1}))$  of the fraction considered in ② are i. i. d., and the first one satisfies

$$\int d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_0), \pi_{\mathbb{T}}(Z_1)) d\mathbb{P} \leq \int d_S(Z_0, Z_1) d\mathbb{P} = \int d_S(1, X_1) d\mathbb{P} < \infty.$$

So, again, by the strong law of large numbers

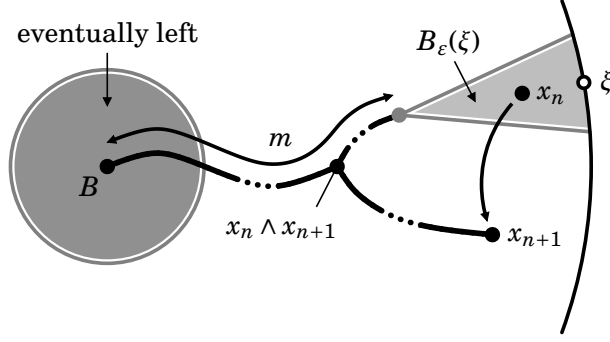
$$\frac{d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_0), \pi_{\mathbb{T}}(Z_1)) + \dots + d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_n), \pi_{\mathbb{T}}(Z_{n+1}))}{n+1} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} \mathbb{E}(d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_0), \pi_{\mathbb{T}}(Z_1))),$$

from where a simple calculation yields ②. □

For the driftless case, the situation is not as easy and in order to show almost sure convergence of the projections  $\pi_{\mathbb{T}}(Z_n)$  to a random end, we restrict ourselves to the non-amenable subcase  $1 < p < q$ .

**Lemma 5.6** *Let  $1 < p < q$ . Assume that  $X_1$  has finite first moment,  $\ln(A_{X_1})$  has finite second moment, and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment. If there is no vertical drift, i. e.  $\delta = 0$ , then the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random end  $\xi \in \partial\mathbb{T}$ .*

*Proof.* Again, we claim that the projections  $\pi_{\mathbb{T}}(Z_n)$  constitute a. s. a regular sequence. But since  $\delta = 0$ , we need to modify the argument from the proof of Lemma 5.5 that showed ①. By assumption,  $G$  is non-amenable and, in particular, the spectral radius  $\varrho(\mu)$  of the random walk  $Z = (Z_0, Z_1, \dots)$  is strictly


 Figure 8: Jumping away from  $B_\varepsilon(\xi)$ .

smaller than 1, see e.g. [Woe00, Corollary 12.5]. This, together with the fact that the random walk is uniformly irreducible yields that

$$\liminf_{n \rightarrow \infty} \frac{d_S(Z_0, Z_n)}{n} > 0.$$

For a proof of this statement, see e.g. [Woe00, Proposition 8.2]. In order to estimate the numerators  $d_S(Z_0, Z_n)$  from above, we apply an auxiliary result: *There are  $\alpha, \beta > 0$  such that for every element  $g \in G$  the inequality  $d_S(1, g) \leq \alpha \cdot |\pi_{\mathbb{T}}(g)|_{\mathbb{T}} + \beta \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g))$  holds.* Let us postpone the proof and record that, using this auxiliary result and Lemma 5.3, we obtain

$$\liminf_{n \rightarrow \infty} \frac{|\pi_{\mathbb{T}}(Z_n)|_{\mathbb{T}}}{n} \geq \frac{1}{\alpha} \cdot \liminf_{n \rightarrow \infty} \left( \frac{d_S(Z_0, Z_n)}{n} - \underbrace{\beta \cdot \frac{d_{\mathbb{H}}(\pi_{\mathbb{H}}(Z_0), \pi_{\mathbb{H}}(Z_n))}{n}}_{\rightarrow 0 \text{ a.s. by Lemma 5.3}} \right) > 0 \text{ a.s.}$$

This is ①. Concerning ②, note that the respective argument from the proof of Lemma 5.5 did not use the assumption that  $\delta \neq 0$ , and therefore does also works for  $\delta = 0$ . So, we know that the projections  $\pi_{\mathbb{T}}(Z_n)$  constitute a.s. a regular sequence, which converges to a random end by the proof of Lemma 5.5.

It remains to show the auxiliary result. In order to do so, we construct a path from 1 to  $g$  in the Cayley graph  $\mathbb{G}$  with at most  $\alpha \cdot |\pi_{\mathbb{T}}(g)|_{\mathbb{T}} + \beta \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g))$  many edges, where the values of  $\alpha$  and  $\beta$  are to be determined uniformly, i.e. not depending on  $g$ . First, we aim to adjust the tree component. Either combinatorially using the defining relation  $ab^pa^{-1} = b^q$  or geometrically using the properties of the Cayley graph  $\mathbb{G}$ , we can find a path from 1 to the coset  $gB$  with at most  $(\lfloor \frac{q}{2} \rfloor + 1) \cdot |\pi_{\mathbb{T}}(g)|_{\mathbb{T}}$  many edges. Let  $h \in gB$  be the endpoint of such a path. Next, recall the notion of discrete hyperbolic plane from Section 2.4. We pick an arbitrary ascending doubly infinite path  $v : \mathbb{Z} \rightarrow G/B$  in the tree  $\mathbb{T}$  that traverses the vertex  $gB$ , consider the discrete hyperbolic plane  $\mathbb{G}_v$ , and take a shortest path from  $h$  to  $g$  in  $\mathbb{G}_v$ . By the proof of Proposition 2.5, its length  $d_{\mathbb{G}_v}(h, g)$  can be estimated from above by  $\kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(h), \pi_{\mathbb{H}}(g))$  with  $\kappa := \max\{\frac{c}{\varepsilon}, \frac{1}{\ell_a}, \frac{1}{\ell_b}\} > 0$ . We may continue this estimate and finally obtain

$$\begin{aligned} d_{\mathbb{G}_v}(h, g) &\leq \kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(h), \pi_{\mathbb{H}}(g)) \\ &\leq \kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(h), \pi_{\mathbb{H}}(1)) + \kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) \\ &\leq \kappa \cdot \max\{\ell_a, \ell_b\} \cdot \left( \left\lfloor \frac{q}{2} \right\rfloor + 1 \right) \cdot |\pi_{\mathbb{T}}(g)|_{\mathbb{T}} + \kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)). \end{aligned}$$

So, the concatenation of the two paths considered above has at most  $\alpha \cdot |\pi_{\mathbb{T}}(g)|_{\mathbb{T}} + \beta \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g))$  many edges with  $\alpha := (1 + \kappa \cdot \max\{\ell_a, \ell_b\}) \cdot (\lfloor \frac{q}{2} \rfloor + 1) > 0$  and  $\beta := \kappa > 0$ .  $\square$



### 5.3 Construction of the Poisson–Furstenberg boundary

Resuming Sections 5.1 and 5.2, we may formulate the following theorem.

**Theorem 5.7 (“convergence theorem”)** *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p < q$ . Suppose that the increment  $X_1$  has finite first moment.*

1. *If the vertical drift is **positive**, i. e.  $\delta > 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a. s. to  $\infty \in \partial\mathbb{H}$  and the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random element  $\xi \in \partial\mathbb{T}$ .*
2. *If the vertical drift is **negative**, i. e.  $\delta < 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a. s. to a random element  $r \in \partial\mathbb{H} \setminus \{\infty\}$  and the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random element  $\xi \in \partial\mathbb{T}$ .*
3. *If there is **no vertical drift**, i. e.  $\delta = 0$ , and, in addition,  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment, which is certainly the case if the increment  $X_1$  has finite  $(2 + \varepsilon)$ -th moment, then the projections  $\pi_{\mathbb{H}}(Z_n)$  have sublinear speed and the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random element  $\xi \in \partial\mathbb{T}$ .*

The boundaries  $\partial\mathbb{H}$  and  $\partial\mathbb{T}$  are endowed with their Borel  $\sigma$ -algebras  $\mathcal{B}_{\partial\mathbb{H}}$  and  $\mathcal{B}_{\partial\mathbb{T}}$ . Under suitable assumptions on the moments, the projections  $\pi_{\mathbb{H}}(Z_n)$  and  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random element in the respective boundary and we may consider the boundary maps  $\text{bnd}_{\partial\mathbb{H}} : \Omega \rightarrow \partial\mathbb{H}$  and  $\text{bnd}_{\partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{T}$ , defined almost everywhere, assigning to a trajectory  $\omega = (x_0, x_1, \dots) \in \Omega$  the limits

$$\text{bnd}_{\partial\mathbb{H}}(\omega) := \lim_{n \rightarrow \infty} \pi_{\mathbb{H}}(x_n) \in \partial\mathbb{H} \quad \text{and} \quad \text{bnd}_{\partial\mathbb{T}}(\omega) := \lim_{n \rightarrow \infty} \pi_{\mathbb{T}}(x_n) \in \partial\mathbb{T}.$$

Even though the boundary maps are only defined almost everywhere, they are measurable in the sense that the preimages of measurable sets are measurable. Given  $\text{bnd}_{\partial\mathbb{H}}$  and  $\text{bnd}_{\partial\mathbb{T}}$ , we may construct their product map  $\text{bnd}_{\partial\mathbb{H} \times \partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{H} \times \partial\mathbb{T}$ . It is measurable with respect to the product  $\sigma$ -algebra  $\mathcal{B}_{\partial\mathbb{H}} \otimes \mathcal{B}_{\partial\mathbb{T}}$ . Since both  $\partial\mathbb{H}$  and  $\partial\mathbb{T}$  are metrisable and separable topological spaces, it is not hard to see that the product  $\sigma$ -algebra  $\mathcal{B}_{\partial\mathbb{H}} \otimes \mathcal{B}_{\partial\mathbb{T}}$  agrees with the Borel  $\sigma$ -algebra  $\mathcal{B}_{\partial\mathbb{H} \times \partial\mathbb{T}}$ , see e. g. [Bil99, Appendix M.10].

In the following, one should keep in mind that  $\text{bnd}_{\partial\mathbb{H}}$ ,  $\text{bnd}_{\partial\mathbb{T}}$ , and  $\text{bnd}_{\partial\mathbb{H} \times \partial\mathbb{T}}$  are only defined if the respective projections  $\pi_{\mathbb{H}}(Z_n)$  and  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random element in the boundary.

**Definition 5.8 (“hitting measures”)** *The three pushforward probability measures  $\nu_{\partial\mathbb{H}} := \text{bnd}_{\partial\mathbb{H}}(\mathbb{P})$ ,  $\nu_{\partial\mathbb{T}} := \text{bnd}_{\partial\mathbb{T}}(\mathbb{P})$ ,  $\nu_{\partial\mathbb{H} \times \partial\mathbb{T}} := \text{bnd}_{\partial\mathbb{H} \times \partial\mathbb{T}}(\mathbb{P})$  on the measurable spaces  $(\partial\mathbb{H}, \mathcal{B}_{\partial\mathbb{H}})$ ,  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}})$ ,  $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{H} \times \partial\mathbb{T}})$  are called the hitting measures. Note that we may again, tacitly, complete the probability spaces with respect to  $\nu_{\partial\mathbb{H}}$ ,  $\nu_{\partial\mathbb{T}}$ ,  $\nu_{\partial\mathbb{H} \times \partial\mathbb{T}}$ .*

Each of the boundaries  $\partial\mathbb{H}$  and  $\partial\mathbb{T}$  is endowed with a left  $G$ -action. The one on  $\partial\mathbb{H}$  is induced by the action  $gz := \pi_{\text{Aff}^+(\mathbb{R})}(g)(z)$  on  $\mathbb{H}$  and the one on  $\partial\mathbb{T}$  is induced by the action  $g(hB) := (gh)B$  on  $\mathbb{T}$ . Let us describe them in more detail. The former is an action by isometries on  $\mathbb{H}$ , and in light of their classification mentioned in Section 2.3, we can also evaluate them at  $\partial\mathbb{H}$ .

For the latter, recall that ends are infinite reduced paths that start in  $B$ . The coordinatewise action on the ends maps every such path  $\xi \in \partial\mathbb{T}$  to some other path that need not start in  $B$  any more. The end  $g\xi \in \partial\mathbb{T}$  is obtained by connecting  $B$  with the initial vertex of this path and reduce the concatenation. This way, it is not hard to see that we can map every  $\xi \in \partial\mathbb{T}$  to an end with an arbitrarily chosen finite

initial piece.<sup>7</sup> In particular, every orbit  $\{g\xi \mid g \in G\}$  is infinite and dense in  $\partial\mathbb{T}$ .

By changing the initial probability measure of the random walk  $Z = (Z_0, Z_1, \dots)$  as in Section 4.2, we obtain stationarity of the measure  $\nu_{\partial\mathbb{T}}$ . More precisely, for every measurable set  $A \subseteq \partial\mathbb{T}$

$$\nu_{\partial\mathbb{T}}(A) = \nu_{\partial\mathbb{T},1}(A) = \sum_{g \in G} \mu(g) \cdot \nu_{\partial\mathbb{T},g}(A) = \sum_{g \in G} \mu(g) \cdot \nu_{\partial\mathbb{T},1}(g^{-1}A) = \sum_{g \in G} \mu(g) \cdot \nu_{\partial\mathbb{T}}(g^{-1}A). \quad (*)$$

The same result holds true for  $\partial\mathbb{H}$  and the product  $\partial\mathbb{H} \times \partial\mathbb{T}$ , which is endowed with the componentwise left  $G$ -action. These observations will be helpful in a moment, when we show that the hitting measures are either Dirac measures or non-atomic. Our proof is based on [Woe89, Lemma 3.4], which is much more general. The original idea for our special case might be older.

**Lemma 5.9** *The hitting measure  $\nu_{\partial\mathbb{T}}$  is non-atomic. Moreover, if  $\delta > 0$ , then the hitting measure  $\nu_{\partial\mathbb{H}}$  is the Dirac measure at  $\infty \in \partial\mathbb{H}$  and, if  $\delta < 0$ , then  $\nu_{\partial\mathbb{H}}$  is again non-atomic.*

*Proof.* Let us first consider the hitting measure  $\nu_{\partial\mathbb{T}}$ . Suppose, there were elements of positive measure. Then, we may choose such an element  $\xi \in \partial\mathbb{T}$  with maximal measure  $a$ . In particular, for every element  $\eta \in \{g\xi \mid g \in G\}$  we know that  $\nu_{\partial\mathbb{T}}(\eta) \leq a$ . We claim that for every  $\eta \in \{g\xi \mid g \in G\}$  the equality  $\nu_{\partial\mathbb{T}}(\eta) = a$  holds. Indeed, let us first suppose that there was an element  $h \in \text{supp}(\mu) \subseteq G$  with  $\nu_{\partial\mathbb{T}}(h^{-1}\xi) < a$ . Then,

$$a = \nu_{\partial\mathbb{T}}(\xi) \stackrel{(*)}{=} \sum_{g \in G} \mu(g) \cdot \nu_{\partial\mathbb{T}}(g^{-1}\xi) = \underbrace{\mu(h) \cdot \nu_{\partial\mathbb{T}}(h^{-1}\xi)}_{< \mu(h) \cdot a} + \underbrace{\sum_{g \in G \setminus \{h\}} \mu(g) \cdot \nu_{\partial\mathbb{T}}(g^{-1}\xi)}_{\leq (1 - \mu(h)) \cdot a}.$$

This is a contradiction. Due to the irreducibility of the random walk,  $\nu_{\partial\mathbb{T}}(h^{-1}\xi) = a$  does not only hold for all  $h \in \text{supp}(\mu) \subseteq G$  but inductively for all  $h \in G$ , which proves our claim. But the orbit  $\{g\xi \mid g \in G\}$  is infinite, so  $1 = \nu_{\partial\mathbb{T}}(\partial\mathbb{T}) \geq |\{g\xi \mid g \in G\}| \cdot a = \infty$ . And this is, again, a contradiction.

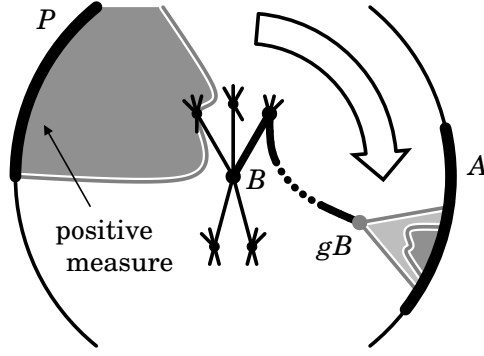
If  $\delta > 0$ , the result that the hitting measure  $\nu_{\partial\mathbb{H}}$  is the Dirac measure at  $\infty \in \partial\mathbb{H}$  is an immediate consequence of Lemma 5.1. On the other hand, if  $\delta < 0$ , then  $\nu_{\partial\mathbb{H}}(\infty) = 0$  by Lemma 5.2. Now, we can repeat the above argument. Suppose, there was an element of positive measure. Then, we may choose such an element  $r \in \partial\mathbb{H} \setminus \{\infty\}$  of maximal measure. Again, all elements in its orbit  $\{gr \mid g \in G\}$  must have the same measure and, since the orbit is infinite, this yields a contradiction.  $\square$

**Lemma 5.10** *The hitting measure  $\nu_{\partial\mathbb{T}}$  has full support, i. e. every non-empty open subset  $A \subseteq \partial\mathbb{T}$  has positive measure. Moreover, if  $\delta < 0$ , the hitting measure  $\nu_{\partial\mathbb{H} \times \partial\mathbb{T}}$  on the Cartesian product  $\partial\mathbb{H} \times \partial\mathbb{T}$  has full support.*

*Proof.* We classify the ends of the tree  $\mathbb{T}$  according to which neighbour of the vertex  $B$  they first traverse. This yields a partition of the space of ends into  $p + q$  open subsets, i. e. the open balls of radius 1. At least one of them must have positive measure, call it  $P \subseteq \partial\mathbb{T}$ .

Now, let  $A \subseteq \partial\mathbb{T}$  be an arbitrary non-empty open subset. In particular, there is a vertex  $gB$ , such that all ends traversing  $gB$  belong to  $A$ . Similarly to the argument given in Footnote 7, either  $gP$  or  $gbP$  is contained in  $A$ , see Figure 9. We may assume w. l. o. g. that  $gP \subseteq A$ . Moreover, due to the irreducibility

<sup>7</sup>Consider such a finite initial piece, i. e. a finite reduced path from  $B$  to some vertex  $gB$ . Given the end  $\xi \in \partial\mathbb{T}$ , we construct its image  $g\xi \in \partial\mathbb{T}$ . It will have the correct finite initial piece unless cancellation takes place. But then, consider the image  $gb\xi \in \partial\mathbb{T}$  instead. Since  $|p| \neq 1$  and  $|q| \neq 1$ , cancellation will take place in at most one of the two cases.


 Figure 9: The hitting measure  $\nu_{\partial T}$  has full support.

of the random walk, there is an  $n \in \mathbb{N}$  such that  $\mu^{(n)}(g) > 0$ , i. e. the probability to reach  $g$  in exactly  $n$  steps is positive. Hence,

$$\nu_{\partial T}(A) = \nu_{\partial T,1}(A) \geq \mu^{(n)}(g) \cdot \nu_{\partial T,g}(A) \geq \mu^{(n)}(g) \cdot \nu_{\partial T,g}(gP) = \mu^{(n)}(g) \cdot \nu_{\partial T,1}(P) = \mu^{(n)}(g) \cdot \nu_{\partial T}(P) > 0.$$

The proof of the second assertion is similar. Since we have to keep track of two components, it is slightly more technical so that we give only a proof sketch. Recall from above the set  $P \subseteq \partial T$ . Given the random variable  $\text{bnd}_{\partial T}$  takes a value in  $P$ , at least one open interval  $(k, k+1) \subseteq \partial \mathbb{H}$  with  $k \in \mathbb{Z}$  will be hit by the random variable  $\text{bnd}_{\partial \mathbb{H}}$  with positive probability, call it  $Q \subseteq \partial \mathbb{H}$ . So, we know that  $\nu_{\partial \mathbb{H} \times \partial T}(Q \times P) > 0$ .

Now, let  $A \subseteq \partial \mathbb{H} \times \partial T$  be an arbitrary non-empty open subset. By definition of the product topology, the open set  $A$  contains a rectangle of open sets  $A_{\partial \mathbb{H}} \subseteq \partial \mathbb{H}$  and  $A_{\partial T} \subseteq \partial T$ . We seek to construct an element  $h \in G$  such that both  $hQ \subseteq A_{\partial \mathbb{H}}$  and  $hP \subseteq A_{\partial T}$ , from where we may finally conclude as above that  $\nu_{\partial \mathbb{H} \times \partial T}(A) > 0$ .

Again, there is a vertex  $gB$  of the tree, such that all ends traversing  $gB$  belong to  $A_{\partial T}$ . Moreover, there are an  $r \in \mathbb{R}$  and an  $\varepsilon > 0$  such that the open interval  $(r, r + \varepsilon)$  is contained in  $A_{\partial \mathbb{H}}$ . Based on this data, we will find an element  $h \in G$  of the form  $h = gb^{k_1}a^{-k_2}b^{k_3}$  with the desired properties.

Let us first look at the tree component. The exponent  $k_1$  is either 0 or 1, whichever ensures that the reduced path from  $B$  to  $gb^{k_1}a^{-1}B$  traverses  $gB$ . Now, let us turn to the hyperbolic component. The image  $gb^{k_1}Q \subseteq \partial \mathbb{H}$  is a bounded open interval. The exponent  $k_2 \in \mathbb{N}$  is chosen in such a way that the length of the image  $gb^{k_1}a^{-k_2}Q \subseteq \partial \mathbb{H}$  is at most  $\frac{\varepsilon}{3}$ . Finally, there is an integer  $k \in \mathbb{Z}$  such that both images  $gb^{k_1}a^{-k_2}b^kQ \subseteq \partial \mathbb{H}$  and  $gb^{k_1}a^{-k_2}b^{k+1}Q \subseteq \partial \mathbb{H}$  are contained in the interval  $(r, r + \varepsilon)$  and therefore both belong to  $A_{\partial \mathbb{H}}$ . The exponent  $k_3$  will be either  $k$  or  $k + 1$ . Let us return to the tree and choose it in such a way that all ends in the image  $gb^{k_1}a^{-k_2}b^{k_3}P$  traverse  $gb^{k_1}a^{-k_2}B$ . Then, by construction, they also traverse the vertex  $gB$  and belong to  $A_{\partial T}$ .  $\square$

**Theorem 5.11 (“identification theorem”)** *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p < q$  and suppose that the increment  $X_1$  has finite first moment.*

1. *If the vertical drift is **positive**, i. e.  $\delta > 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial T, \mathcal{B}_{\partial T}, \nu_{\partial T})$  endowed with the boundary map  $\text{bnd}_{\partial T} : \Omega \rightarrow \partial T$ .*

2. If the vertical drift is **negative**, i. e.  $\delta < 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{H} \times \partial\mathbb{T}}, \nu_{\partial\mathbb{H} \times \partial\mathbb{T}})$  endowed with the boundary map  $\text{bnd}_{\partial\mathbb{H} \times \partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{H} \times \partial\mathbb{T}$ .
3. If there is **no vertical drift**, i. e.  $\delta = 0$ , and, in addition,  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment, which is certainly the case if the increment  $X_1$  has finite  $(2 + \varepsilon)$ -th moment, then the Poisson–Furstenberg boundary is again isomorphic to  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  endowed with the boundary map  $\text{bnd}_{\partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{T}$ .

**Remark 5.12** In the case of negative vertical drift, we could also have used  $\mathbb{R} \times \partial\mathbb{T}$  instead of  $\partial\mathbb{H} \times \partial\mathbb{T}$ .

*Proof.* As already mentioned, we seek to apply the strip criterion, see Theorem 4.3. By Lemma 4.5, the probability measure  $\mu$  driving the random walk has finite entropy. Moreover, it is not hard to see that  $\partial\mathbb{H}$  and  $\partial\mathbb{T}$  are Polish spaces, and so is their product  $\partial\mathbb{H} \times \partial\mathbb{T}$ . Therefore, by Remark 4.1, the probability spaces  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  and  $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{H} \times \partial\mathbb{T}}, \nu_{\partial\mathbb{H} \times \partial\mathbb{T}})$  are Lebesgue–Rohlin spaces. They are endowed with a left  $G$ -action and boundary maps  $\text{bnd}_{\partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{T}$  and  $\text{bnd}_{\partial\mathbb{H} \times \partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{H} \times \partial\mathbb{T}$ , defined almost everywhere. In order to show that they are  $\mu$ -boundaries, we have to ensure that the boundary maps are ① measurable, ②  $\sim$ -invariant, and ③  $G$ -equivariant. But all three properties are immediate by construction, compare also [Kai00, end of §1.5].

If the vertical drift is negative, i. e.  $\delta < 0$ , let us take the  $\mu$ -boundary  $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{H} \times \partial\mathbb{T}}, \nu_{\partial\mathbb{H} \times \partial\mathbb{T}})$  and the  $\check{\mu}$ -boundary  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \check{\nu}_{\partial\mathbb{T}})$ . Here,  $\check{\nu}_{\partial\mathbb{T}}$  denotes the hitting measure of the pointwise projection of the random walk  $\check{Z} = (\check{Z}_0, \check{Z}_1, \dots)$  driven by the reflected probability measure  $\check{\mu}$  to the Bass–Serre tree  $\mathbb{T}$ .

Next, we need to define gauges and strips. Let  $S := \{a, b\} \subseteq G$  be the standard generating set and define gauges  $\mathcal{G}_k := \{g \in G \mid d_S(1, g) \leq k\}$ . In other words, the gauges exhaust the group  $G$  with balls centred at the identity  $1 \in G$ , and the gauge function  $|\cdot| = |\cdot|_{\mathcal{G}}$  is nothing but the distance to 1 with respect to the word metric  $d_S$ .

By Lemma 5.9, we know that  $\check{\nu}_{\partial\mathbb{T}} \otimes \nu_{\partial\mathbb{H} \times \partial\mathbb{T}}$ -almost every pair of points  $(\xi_-, (r_+, \xi_+)) \in \partial\mathbb{T} \times (\partial\mathbb{H} \times \partial\mathbb{T})$  has distinct ends  $\xi_-, \xi_+ \in \partial\mathbb{T}$  and a boundary value  $r_+ \in \mathbb{R}$ . In this situation, we may connect  $\xi_-$  and  $\xi_+$  by a unique doubly infinite reduced path  $v : \mathbb{Z} \rightarrow \mathbb{T}$  and define the strip  $\mathcal{S}(\xi_-, (r_+, \xi_+))$  as follows. It consists of all  $g \in G$  that are contained in the full  $\pi_{\mathbb{T}}$ -preimage of  $v(\mathbb{Z})$ , i. e. their image  $\pi_{\mathbb{T}}(g)$  is traversed by  $v$ , and have the property that the real part  $\text{Re}(\pi_{\mathbb{H}}(g))$  has minimal distance to  $r_+ \in \mathbb{R}$  among all real parts  $\text{Re}(\pi_{\mathbb{H}}(h))$  with  $h \in gB$ , see the left-hand side of Figure 10. To all remaining pairs we assign the whole of  $G$  as a strip. This way, the map  $\mathcal{S}$  becomes measurable and  $G$ -equivariant. By Lemma 5.10, a random strip contains the identity element  $1 \in G$  with positive probability, i. e. the map  $\mathcal{S}$  satisfies the inequality of Remark 4.4. So, it suffices to verify the following convergence for an arbitrary pair  $(\xi_-, (r_+, \xi_+)) \in \partial\mathbb{T} \times (\partial\mathbb{H} \times \partial\mathbb{T})$  with distinct ends  $\xi_-, \xi_+ \in \partial\mathbb{T}$  and  $r_+ \in \mathbb{R}$ ,

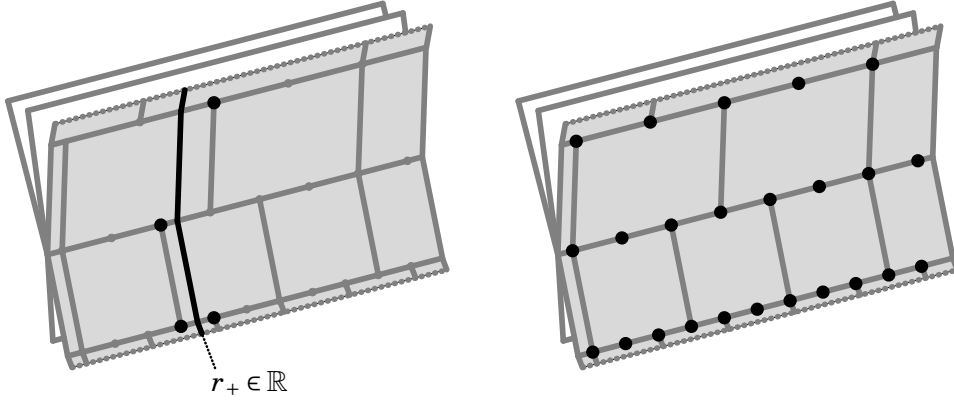
$$\frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(\xi_-, (r_+, \xi_+)) \cap \mathcal{G}_{|Z_n|})) \xrightarrow[n \rightarrow \infty]{\text{a. s.}} 0.$$

But the strip  $\mathcal{S}(\xi_-, (r_+, \xi_+))$  intersects the gauge  $\mathcal{G}_{|Z_n|}$  in at most  $2 \cdot |Z_n| + 1$  many cosets of the form  $G/B$ , and each of them contains at most two elements of the strip. Therefore,

$$\frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(\xi_-, (r_+, \xi_+)) \cap \mathcal{G}_{|Z_n|})) \leq \frac{\ln((2 \cdot |Z_n| + 1) \cdot 2)}{n} = \frac{\ln((2 \cdot d_S(1, Z_n) + 1) \cdot 2)}{n} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} 0.$$

In the final step of the above calculation, we use again that the increments  $X_1$  have finite first moment. Indeed,

$$\frac{1}{n} \cdot d_S(1, Z_n) = \frac{1}{n} \cdot d_S(1, X_1 \cdot \dots \cdot X_n) \leq \frac{1}{n} \cdot \sum_{k=1}^n d_S(1, X_k) \xrightarrow[n \rightarrow \infty]{\text{a. s.}} \mathbb{E}(d_S(1, X_1)),$$


 Figure 10: Strips for the cases  $\delta \neq 0$  (left) and  $\delta = 0$  (right).

from where we may first conclude that the sequence  $\frac{1}{n} \cdot d_S(1, Z_n)$  is a. s. bounded and second that the sequence  $\frac{1}{n} \cdot \ln((2 \cdot d_S(1, Z_n) + 1) \cdot 2)$  converges a. s. to 0.

So, we can finally apply the strip criterion and obtain that  $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{H} \times \partial\mathbb{T}}, \nu_{\partial\mathbb{H} \times \partial\mathbb{T}})$  is isomorphic to the Poisson–Furstenberg boundary. Vice versa, if the vertical drift is positive, i. e.  $\delta > 0$ , the same argument yields that  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  is isomorphic to the Poisson–Furstenberg boundary.

It remains to consider the driftless case, i. e.  $\delta = 0$ . Then, both  $\mu$  and  $\check{\mu}$  are driftless and there is no natural candidate for a real number that determines the horizontal position of the strip. But the fact that the projections  $\pi_{\mathbb{H}}(Z_n)$  have sublinear speed allows us to solve this issue. More precisely, take the  $\mu$ -boundary  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  and the  $\check{\mu}$ -boundary  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \check{\nu}_{\partial\mathbb{T}})$ . Now, define gauges

$$\mathcal{G}_k := \{g \in G \mid d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) \leq k \text{ and } d_{\mathbb{T}}(\pi_{\mathbb{T}}(1), \pi_{\mathbb{T}}(g)) \leq k^2\}.$$

Again, we know that  $\check{\nu}_{\partial\mathbb{T}} \otimes \nu_{\partial\mathbb{T}}$ -almost every pair of points  $(\xi_-, \xi_+) \in \partial\mathbb{T} \times \partial\mathbb{T}$  has distinct ends  $\xi_-, \xi_+ \in \partial\mathbb{T}$ , which we may connect by a unique doubly infinite reduced path  $v : \mathbb{Z} \rightarrow \mathbb{T}$ . Let  $\mathcal{S}(\xi_-, \xi_+)$  be the full  $\pi_{\mathbb{T}}$ -preimage of  $v(\mathbb{Z})$ , i. e. the set of all group elements  $g \in G$  such that the image  $\pi_{\mathbb{T}}(g)$  is traversed by  $v$ , see the right-hand side of Figure 10. Again, to all remaining pairs we assign the whole of  $G$  as a strip. This way, the map  $\mathcal{S}$  becomes measurable,  $G$ -equivariant, and satisfies the inequality of Remark 4.4. Now, pick an arbitrary pair  $(\xi_-, \xi_+) \in \partial\mathbb{T} \times \partial\mathbb{T}$  with distinct ends  $\xi_-, \xi_+ \in \partial\mathbb{T}$ . We claim that

$$\frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(\xi_-, \xi_+) \cap \mathcal{G}_{|Z_n|})) \leq \frac{\ln((2 \cdot |Z_n|^2 + 1) \cdot \exp(|Z_n| + 2))}{n} = \underbrace{\frac{\ln(2 \cdot |Z_n|^2 + 1)}{n}}_{\textcircled{1}} + \underbrace{\frac{|Z_n| + 2}{n}}_{\textcircled{2}}.$$

Indeed, the inequality holds for a similar reason as above; the strip  $\mathcal{S}(\xi_-, \xi_+)$  intersects the gauge  $\mathcal{G}_{|Z_n|}$  in at most  $2 \cdot |Z_n|^2 + 1$  many cosets of the form  $G/B$ . Slightly more involved is the observation that each of them contains at most  $\exp(|Z_n| + 2)$  many elements of the gauge. Fix a coset  $gB$ . The projections  $\pi_{\mathbb{H}}(h)$  of the elements  $h \in gB$  are located on the horizontal line  $L \subseteq \mathbb{H}$  with imaginary part  $y := \text{Im}(\pi_{\mathbb{H}}(g))$ . One necessary condition for such an element  $h \in gB$  to be contained in the gauge  $\mathcal{G}_{|Z_n|}$  is that the projection  $\pi_{\mathbb{H}}(h)$  is contained in the closed disc  $D := \{z \in \mathbb{H} \mid d_{\mathbb{H}}(i, z) \leq |Z_n|\} \subseteq \mathbb{H}$ . If  $L \cap D$  is empty, then the coset  $gB$  does not contain any element of the gauge and we are done. Otherwise, there is a unique  $x \in \mathbb{R}$  with  $x \geq 0$  such that  $L \cap D$  is the horizontal line between  $z_1 := -x + iy$  and  $z_2 := x + iy$ , see Figure 11. The projections  $\pi_{\mathbb{H}}(h)$  with  $h \in gB$  have the property that the real parts  $\text{Re}(\pi_{\mathbb{H}}(h))$  and  $\text{Re}(\pi_{\mathbb{H}}(hb))$  differ exactly by  $y$ .

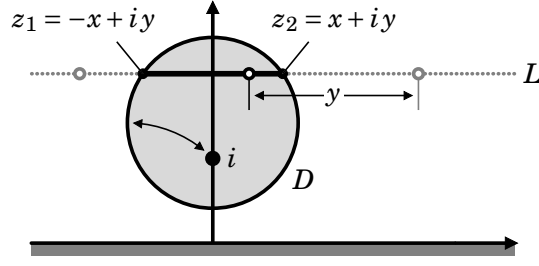


Figure 11: The horizontal line  $L$ , the closed disc  $D$ , and their intersection  $L \cap D$ .

So, the horizontal line  $L \cap D$  contains at most  $1 + \frac{2x}{y}$  many of them. Let us now estimate  $1 + \frac{2x}{y}$  in terms of  $|Z_n|$ . Since  $z_1$  and  $z_2$  are both contained in  $D$ , their distance is at most  $2 \cdot |Z_n|$ . Therefore,

$$2 \cdot |Z_n| \geq d_{\mathbb{H}}(z_1, z_2) = \operatorname{arcosh} \left( 1 + \frac{|z_2 - z_1|^2}{2 \operatorname{Im}(z_1) \operatorname{Im}(z_2)} \right) = \operatorname{arcosh} \left( 1 + \frac{2x^2}{y^2} \right) \geq \ln \left( 1 + \frac{2x^2}{y^2} \right).$$

And, in particular,

$$\begin{aligned} \exp(2 \cdot |Z_n|) &\geq 1 + \frac{2x^2}{y^2}, \implies \exp(2 \cdot |Z_n|) > \frac{2x^2}{y^2}, \implies \exp(2 \cdot |Z_n| + \ln(2)) > \frac{4x^2}{y^2}, \\ &\implies \exp \left( |Z_n| + \frac{1}{2} \cdot \ln(2) \right) > \frac{2x}{y}, \implies \exp(|Z_n| + 2) > 1 + \frac{2x}{y}. \end{aligned}$$

So, the coset  $gB$  contains strictly fewer than  $\exp(|Z_n| + 2)$  elements of the gauge. We will now show that both summands ① and ② converge a. s. to 0, which will complete the proof. Let us first observe that

$$|Z_n| - 1 \leq \max \left\{ d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_{\mathbb{T}}(\pi_{\mathbb{T}}(1), \pi_{\mathbb{T}}(Z_n))} \right\} \leq \max \left\{ d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_{\mathbb{S}}(1, Z_n)} \right\}. \quad (*)$$

Concerning ①, we deduce from  $(*)$  and the proof of Proposition 2.5 that  $|Z_n| \leq \max\{\ell_a, \ell_b, 1\} \cdot d_{\mathbb{S}}(1, Z_n) + 1$ , and finally obtain by the same argument as above

$$\textcircled{1} = \frac{\ln(2 \cdot |Z_n|^2 + 1)}{n} \leq \frac{\ln(2 \cdot (\max\{\ell_a, \ell_b, 1\} \cdot d_{\mathbb{S}}(1, Z_n) + 1)^2 + 1)}{n} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} 0.$$

On the other hand, concerning ②, we apply  $(*)$  and Lemma 5.3 to obtain

$$\textcircled{2} = \frac{|Z_n| + 2}{n} \leq \frac{\max \left\{ d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_{\mathbb{S}}(1, Z_n)} \right\} + 3}{n} \xrightarrow[n \rightarrow \infty]{\text{a. s.}} 0.$$

□

## Appendix: The remaining non-amenable cases

Recall from Section 2.1 that a Baumslag–Solitar group  $\operatorname{BS}(p, q)$  is non-amenable if and only if neither  $|p| = 1$  nor  $|q| = 1$ . Until now, we have only identified the Poisson–Furstenberg boundary for random walks on non-amenable Baumslag–Solitar groups  $\operatorname{BS}(p, q)$  with  $1 < p < q$ . Replacing one of the two generators by its inverse, it is easy to see that

$$\operatorname{BS}(p, q) \cong \operatorname{BS}(q, p) \quad \text{and} \quad \operatorname{BS}(p, q) \cong \operatorname{BS}(-p, -q).$$

So, in order to investigate the remaining non-amenable Baumslag–Solitar groups, the only cases that we have to consider are  $1 < p < -q$  and  $1 < p = |q|$ . In this appendix, we shall review our methods from the previous sections and explain how to adjust the arguments in order to obtain similar results for these cases.

### A.1 Action by suitable isometries on the hyperbolic plane

Let us first assume that  $G = \text{BS}(p, q)$  with  $1 < p < -q$ . In order to define the projection  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$  back in Section 2.3, we considered the map  $\pi_{\text{Aff}^+(\mathbb{R})} : \{a, b\} \rightarrow \text{Aff}^+(\mathbb{R})$  given by  $\pi_{\text{Aff}^+(\mathbb{R})}(a) := (x \mapsto \frac{q}{p} \cdot x)$  and  $\pi_{\text{Aff}^+(\mathbb{R})}(b) := (x \mapsto x + 1)$ , and extended it to a homomorphism.

Now, we are assuming that  $1 < p < -q$ , in which case the transformation  $x \mapsto \frac{q}{p} \cdot x$  is not orientation preserving any more. If we replaced  $q$  by  $|q|$  in the definition, then  $\pi_{\text{Aff}^+(\mathbb{R})}(a) := (x \mapsto \frac{|q|}{p} \cdot x)$  would be orientation preserving but  $\pi_{\text{Aff}^+(\mathbb{R})}(a) \circ \pi_{\text{Aff}^+(\mathbb{R})}(a)^p \circ \pi_{\text{Aff}^+(\mathbb{R})}(a)^{-1} \neq \pi_{\text{Aff}^+(\mathbb{R})}(b)^q$ , whence we could not apply von Dyck’s theorem any more. So, we have to change the approach.

Let  $M$  be the set of all maps  $\varphi : \mathbb{C} \rightarrow \mathbb{C}$  either of the form  $\varphi(z) = \alpha z + \beta$  or of the form  $\varphi(z) = \alpha \cdot (-\bar{z}) + \beta$  with  $\alpha, \beta \in \mathbb{R}$  and  $\alpha > 0$ . This set endowed with the composition forms again a group. Consider the map  $\pi_M : \{a, b\} \rightarrow M$  given by  $\pi_M(a) := (z \mapsto \frac{|q|}{p} \cdot (-\bar{z}))$  and  $\pi_M(b) := (z \mapsto z + 1)$ . With this map, it is possible to apply von Dyck’s theorem and to extend it uniquely to a group homomorphism  $\pi_M : G \rightarrow M$ . Finally, as in the case of  $\text{Aff}^+(\mathbb{R})$ , every  $\varphi \in M$  can be thought of as an isometry of  $\mathbb{H}$ . So, we may again consider the projection  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$  given by  $\pi_{\mathbb{H}}(g) := \pi_M(g)(i)$ . The following lemma illustrates this definition.

**Lemma A.1** *For every  $g \in G$  the point  $\pi_{\mathbb{H}}(ga) \in \mathbb{H}$  is above the point  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ ; the two points have the same real part and their distance is  $\ell_a := \ln\left(\frac{|q|}{p}\right)$ . But for every  $g \in G$  the point  $\pi_{\mathbb{H}}(gb) \in \mathbb{H}$  is either right or left from the point  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$  depending on whether the level  $\lambda(g)$  is even or odd; in any case, the two points have the same imaginary part and their distance is  $\ell_b := \ln\left(\frac{3+\sqrt{5}}{2}\right)$ .*

*Proof sketch.* The proof is similar to the one of Lemma 2.3. So, we only discuss the differences. Let us consider the two points  $\pi_{\mathbb{H}}(1) \in \mathbb{H}$  and  $\pi_{\mathbb{H}}(b) \in \mathbb{H}$ . If  $\pi_M(g)$  is of the form  $z \mapsto \alpha z + \beta$ , then it is again the composition of a dilation  $z \mapsto \alpha z$  and a translation  $z \mapsto z + \beta$ , whence the relative position of the two points is preserved. On the other hand, if  $\pi_M(g)$  is of the form  $z \mapsto \alpha \cdot (-\bar{z}) + \beta$ , then it is the composition of a reflection at the imaginary axis  $z \mapsto -\bar{z}$ , a dilation  $z \mapsto \alpha z$ , and a translation  $z \mapsto z + \beta$ , in which case the relative position of the two points is still preserved with the exception that right and left are switched.

In order to decide whether  $\pi_M(g)$  is of the first or the second form, we can write the element  $g \in G$  as a product over  $a^{\pm 1}$  and  $b^{\pm 1}$ . Since  $\pi_M : G \rightarrow M$  is a homomorphism, the image  $\pi_M(g)$  can be written as the respective product over  $\pi_M(a^{\pm 1})$  and  $\pi_M(b^{\pm 1})$ . But each occurrence of  $\pi_M(a^{\pm 1})$  yields one reflection. So, the image  $\pi_M(g)$  is of the first form if and only if the number of occurrences of  $\pi_M(a^{\pm 1})$  is even, which is the case if and only if  $\lambda(g)$  is even.  $\square$

Using this projection  $\pi_{\mathbb{H}} : G \rightarrow \mathbb{H}$  and, of course, replacing  $q$  by  $|q|$  wherever it is necessary, we can repeat the arguments from the previous sections. The definition of the tree  $\mathbb{T}$  and the level functions  $\lambda$  and  $\tilde{\lambda}$ , including Lemma 2.2, as well as the definition of the discrete hyperbolic plane  $\mathbb{G}_v$ , including Proposition 2.5, can be adapted. Recall that, in Section 3.3, we considered the imaginary and real parts



of  $\pi_{\mathbb{H}}(g) \in \mathbb{H}$  separately, and introduced the shorthand notation  $A_g := \text{Im}(\pi_{\mathbb{H}}(g))$  and  $B_g := \text{Re}(\pi_{\mathbb{H}}(g))$ . Let us highlight that we have  $\ln(A_g) = \ln\left(\frac{|q|}{p}\right) \cdot \lambda(g)$ , which allows us to adapt the proof of Lemma 3.4.

In order to identify the Poisson–Furstenberg boundary geometrically, we have to ensure convergence to the boundaries  $\partial\mathbb{H}$  and  $\partial\mathbb{T}$ . Let us first consider the boundary  $\partial\mathbb{H}$ . The proof of Lemma 5.1 for  $\delta > 0$  can be adapted. The proof of Lemma 5.2 for  $\delta < 0$ , in turn, deserves a bit of work. We have to show that the real parts  $B_{Z_n}$  converge a. s. to a random element  $r \in \partial\mathbb{H} \setminus \{\infty\}$ . In the original proof, we observed that  $A_{Z_n} = A_{X_1} \cdots A_{X_n}$  and  $B_{Z_n} = \sum_{k=1}^n C_k$  with  $C_k := A_{X_1} \cdots A_{X_{k-1}} \cdot B_{X_k}$ . While the first formula remains true, the second one does not. We are now in a situation where not only the scaling but also the direction of the next horizontal increment depends on the current level. However, instead of the above, we obtain that  $C_k := \varepsilon_{X_1} \cdot A_{X_1} \cdots \varepsilon_{X_{k-1}} \cdot A_{X_{k-1}} \cdot B_{X_k}$  with  $\varepsilon_g := 1$  if  $\lambda(g)$  is even and  $\varepsilon_g := -1$  if  $\lambda(g)$  is odd. This observation allows us to apply Cauchy’s root test precisely as in the proof of Lemma 5.2. For the same reason, namely because all the estimates are not in terms of the actual horizontal increments but of their absolute values, the proofs of Lemmas 5.3 and 5.4 for  $\delta = 0$  can be adapted. The same holds, concerning the boundary  $\partial\mathbb{T}$ , for the proofs of Lemma 5.5 for  $\delta \neq 0$  and Lemma 5.6 for  $\delta = 0$ . From these observations, we may deduce the following results.

**Theorem A.2 (“convergence theorem” for  $1 < p < -q$ )** *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p < -q$  and suppose that the increment  $X_1$  has finite first moment.*

1. *If the vertical drift is **positive**, i. e.  $\delta > 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a. s. to  $\infty \in \partial\mathbb{H}$  and the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random element  $\xi \in \partial\mathbb{T}$ .*
2. *If the vertical drift is **negative**, i. e.  $\delta < 0$ , then the projections  $\pi_{\mathbb{H}}(Z_n)$  converge a. s. to a random element  $r \in \partial\mathbb{H} \setminus \{\infty\}$  and the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random element  $\xi \in \partial\mathbb{T}$ .*
3. *If there is **no vertical drift**, i. e.  $\delta = 0$ , and, in addition,  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment, which is certainly the case if the increment  $X_1$  has finite  $(2 + \varepsilon)$ -th moment, then the projections  $\pi_{\mathbb{H}}(Z_n)$  have sublinear speed and the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random element  $\xi \in \partial\mathbb{T}$ .*

**Theorem A.3 (“identification theorem” for  $1 < p < -q$ )** *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p < -q$  and suppose that the increment  $X_1$  has finite first moment.*

1. *If the vertical drift is **positive**, i. e.  $\delta > 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  endowed with the boundary map  $\text{bnd}_{\partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{T}$ .*
2. *If the vertical drift is **negative**, i. e.  $\delta < 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{H} \times \partial\mathbb{T}}, \nu_{\partial\mathbb{H} \times \partial\mathbb{T}})$  endowed with the boundary map  $\text{bnd}_{\partial\mathbb{H} \times \partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{H} \times \partial\mathbb{T}$ .*
3. *If there is **no vertical drift**, i. e.  $\delta = 0$ , and, in addition,  $\ln(A_{X_1})$  has finite second moment and there is an  $\varepsilon > 0$  such that  $\ln(1 + |B_{X_1}|)$  has finite  $(2 + \varepsilon)$ -th moment, which is certainly the case if the increment  $X_1$  has finite  $(2 + \varepsilon)$ -th moment, then the Poisson–Furstenberg boundary is again isomorphic to  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  endowed with the boundary map  $\text{bnd}_{\partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{T}$ .*

## A.2 Action by isometries on the Euclidean plane

Let us now assume that  $G = \text{BS}(p, q)$  with  $1 < p = |q|$ . This situation differs fundamentally from the ones discussed so far because the bricks introduced in the proof of Proposition 2.5 would now have equally many  $\mathbb{H}$ -vertices on their upper and lower level. Therefore, we use the Euclidean plane  $\mathbb{R}^2$  instead of the hyperbolic plane  $\mathbb{H}$ . In order to construct a projection  $\pi_{\mathbb{R}^2} : G \rightarrow \mathbb{R}^2$ , let  $M := \text{Isom}(\mathbb{R}^2)$  and consider the map  $\pi_M : \{a, b\} \rightarrow M$  given by

$$\pi_M(a) := \begin{cases} ((x, y) \mapsto (x, y + 1)) & \text{if } q > 0 \\ ((x, y) \mapsto (-x, y + 1)) & \text{if } q < 0 \end{cases} \quad \text{and} \quad \pi_M(b) := ((x, y) \mapsto (x + 1, y)).$$

In both cases, i. e.  $q > 0$  and  $q < 0$ , it is possible to apply von Dyck's theorem and to extend the map uniquely to a group homomorphism  $\pi_{\mathbb{R}^2} : G \rightarrow M$ . Now, we may consider the projection  $\pi_{\mathbb{R}^2} : G \rightarrow \mathbb{R}^2$  given by  $\pi_{\mathbb{R}^2}(g) := \pi_M(g)(0, 0)$ .

The definition of the tree  $\mathbb{T}$  and the level functions  $\lambda$  and  $\tilde{\lambda}$ , including Lemma 2.2, remain the same. But instead of the discrete hyperbolic plane, we now obtain a discrete Euclidean plane  $\mathbb{G}_v$ . The proof of Proposition 2.5 can be adapted to the new situation and shows that the graph  $\mathbb{G}_v$  endowed with the graph distance  $d_{\mathbb{G}_v}$  is quasi-isometric, and even bi-Lipschitz, to the Euclidean plane  $\mathbb{R}^2$  endowed with the standard metric  $d_{\mathbb{R}^2}$ .

We aim to show that, as soon as the projections converge to a random element in  $\partial\mathbb{T}$ , independently of the vertical drift, the Poisson–Furstenberg boundary is isomorphic to  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$ . In particular, we do not need to introduce any boundary to capture the behaviour of the projections  $\pi_{\mathbb{R}^2}(Z_n)$ . Concerning the projections  $\pi_{\mathbb{T}}(Z_n)$ , we distinguish between two cases. If the vertical drift is different from 0, i. e.  $\delta \neq 0$ , then the proof of Lemma 5.5 can be adapted and we obtain that the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random end  $\xi \in \partial\mathbb{T}$ . But if there is no vertical drift, i. e.  $\delta = 0$ , then the proof of Lemma 5.6 cannot be adapted because it was based on the fact that the projections  $\pi_{\mathbb{H}}(Z_n)$  had sublinear speed; and the projections  $\pi_{\mathbb{R}^2}(Z_n)$  do not need to have sublinear speed any more. In this situation, the following lemma may be used instead of Lemma 5.6.

**Lemma A.4** *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on an arbitrary non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$ . If the increment  $X_1$  has finite support, then the projections  $\pi_{\mathbb{T}}(Z_n)$  converge a. s. to a random end  $\xi \in \partial\mathbb{T}$ .*

*Proof sketch.* Recall from the beginning of Section 5.2 that the projections  $\pi_{\mathbb{T}}(Z_n)$  do not need to satisfy the Markov property. Despite of this, we first show that they leave a. s. every finite ball with centre  $B$ , i. e. every set of vertices of the form  $\{x \in G/B \mid d_{\mathbb{T}}(B, x) < r\}$ . Suppose they did not. Then, there is a ball such that the probability to visit this ball infinitely often is strictly positive. Now, it is not hard to see that the probability to visit the centre of this ball infinitely often is also strictly positive. In other words,

$$\mathbb{P}(\{\omega \in \Omega \mid \exists \text{ infinitely many } n \in \mathbb{N} \text{ such that } \pi_{\mathbb{T}}(Z_n(\omega)) = B\}) > 0.$$

But for every  $\omega \in \Omega$  we obtain

$$\begin{aligned} \pi_{\mathbb{T}}(Z_n(\omega)) = B &\iff Z_n(\omega) \cdot B = B \iff Z_n(\omega) \in B \iff Z_n(\omega)^{-1} \in B \\ &\iff Z_n(\omega)^{-1} \cdot B = B \iff X_n(\omega)^{-1} \cdot \dots \cdot X_1(\omega)^{-1} \cdot B = B. \end{aligned}$$

## REFERENCES

Therefore,  $\mathbb{P}(\{\omega \in \Omega \mid \exists \text{ infinitely many } n \in \mathbb{N} \text{ such that } X_n(\omega)^{-1} \cdot \dots \cdot X_1(\omega)^{-1} \cdot B = B\}) > 0$ . We may read the random sequence  $L = (L_0, L_1, \dots)$  with  $L_n := X_n^{-1} \cdot \dots \cdot X_1^{-1} \cdot B$  as the left random walk on the left homogeneous space  $G/B$  driven by the reflected probability measure  $\check{\mu}$ . By the above,  $L$  is recurrent. But since the group  $G$  is non-amenable and the subgroup  $B = \langle b \rangle \leq G$  is amenable, one can show that the homogeneous space  $G/B$  is non-amenable, see [Sch81] referring to [Eym72, §1.3, Exemple 2b]. Therefore, by [Sch81, Satz II.9],  $L$  must be transient, which is a contradiction.

Now, we can repeat the argument given in the proof of Lemma 5.5 to show that for almost every trajectory  $\omega \in \Omega$  there is an end  $\xi(\omega) \in \partial\mathbb{T}$  such that for every  $\varepsilon > 0$  the projections  $\pi_{\mathbb{T}}(Z_n(\omega))$  visit the open ball  $B_\varepsilon(\xi(\omega)) := \{x \in \widehat{\mathbb{T}} \mid d_{\widehat{\mathbb{T}}}(\xi(\omega), x) < \varepsilon\}$  infinitely often. Moreover, we know that for almost every trajectory  $\omega \in \Omega$  both ① the projections  $\pi_{\mathbb{T}}(Z_n(\omega))$  leave every finite ball with centre  $B$  and ② the graph distance of any two subsequent projections  $\pi_{\mathbb{T}}(Z_n(\omega))$  and  $\pi_{\mathbb{T}}(Z_{n+1}(\omega))$  is bounded by a constant. Hence, for every  $\varepsilon > 0$  the projections  $\pi_{\mathbb{T}}(Z_n(\omega))$  must eventually remain in the open ball  $B_\varepsilon(\xi(\omega))$ . In other words, they converge to  $\xi(\omega) \in \partial\mathbb{T}$ .  $\square$

Now, we can show as in Lemmas 5.9 and 5.10 that the hitting measure  $\nu_{\mathbb{T}}$  is again non-atomic and has full support. Moreover, we obtain the following version of the identification theorem.

**Theorem A.5 (“identification theorem” for  $1 < p = |q|$ )** *Let  $Z = (Z_0, Z_1, \dots)$  be a random walk on a non-amenable Baumslag–Solitar group  $G = \text{BS}(p, q)$  with  $1 < p = |q|$  and suppose that the increment  $X_1$  has finite first moment.*

1. *If the **vertical drift** is different from 0, i.e.  $\delta \neq 0$ , then the Poisson–Furstenberg boundary is isomorphic to  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  endowed with the boundary map  $\text{bnd}_{\partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{T}$ .*
2. *If there is **no vertical drift**, i.e.  $\delta = 0$ , and the increment  $X_1$  has not just finite first moment but also finite support, then the Poisson–Furstenberg boundary is again isomorphic to  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  endowed with the boundary map  $\text{bnd}_{\partial\mathbb{T}} : \Omega \rightarrow \partial\mathbb{T}$ .*

*Proof sketch.* Again, we apply the strip criterion. As in the original proof of the identification theorem, we take the  $\mu$ -boundary  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$  and the  $\check{\mu}$ -boundary  $(\partial\mathbb{T}, \mathcal{B}_{\partial\mathbb{T}}, \check{\nu}_{\partial\mathbb{T}})$ . Next, we define gauges

$$\mathcal{G}_k := \{g \in G \mid d_{\mathbb{R}^2}(\pi_{\mathbb{R}^2}(1), \pi_{\mathbb{R}^2}(g)) \leq k \text{ and } d_{\mathbb{T}}(\pi_{\mathbb{T}}(1), \pi_{\mathbb{T}}(g)) \leq k\}.$$

We know that  $\check{\nu}_{\partial\mathbb{T}} \otimes \nu_{\partial\mathbb{T}}$ -almost every pair of points  $(\xi_-, \xi_+) \in \partial\mathbb{T} \times \partial\mathbb{T}$  has distinct ends  $\xi_-, \xi_+ \in \partial\mathbb{T}$ , which we may connect by a unique doubly infinite reduced path  $v : \mathbb{Z} \rightarrow \mathbb{T}$ . Let  $\mathcal{S}(\xi_-, \xi_+)$  be the full  $\pi_{\mathbb{T}}$ -preimage of  $v(\mathbb{Z})$ . To all remaining pairs we assign the whole of  $G$  as a strip. This way, the map  $\mathcal{S}$  becomes measurable,  $G$ -equivariant, and satisfies the inequality of Remark 4.4. Now, pick an arbitrary pair  $(\xi_-, \xi_+) \in \partial\mathbb{T} \times \partial\mathbb{T}$  with distinct ends  $\xi_-, \xi_+ \in \partial\mathbb{T}$ . We claim that

$$\frac{1}{n} \cdot \ln(\text{card}(\mathcal{S}(\xi_-, \xi_+) \cap \mathcal{G}_{|Z_n|})) \leq \frac{\ln((2 \cdot |Z_n| + 1) \cdot (2 \cdot |Z_n| + 1))}{n}.$$

Indeed, the inequality is easy to see. The strip  $\mathcal{S}(\xi_-, \xi_+)$  intersects the gauge  $\mathcal{G}_{|Z_n|}$  in at most  $2 \cdot |Z_n| + 1$  many cosets of the form  $G/B$ , and each of them contains at most  $2 \cdot |Z_n| + 1$  many elements of the gauge. Now, it suffices to consider the standard generating set  $S := \{a, b\} \subseteq G$  and to observe that  $|Z_n| \leq d_S(1, Z_n) + 1$ . Then, using the fact that  $\frac{1}{n} \cdot d_S(1, Z_n)$  is a.s. bounded, we may conclude that

$$\dots = \frac{\ln((2 \cdot |Z_n| + 1)^2)}{n} \leq \frac{\ln((2 \cdot d_S(1, Z_n) + 3)^2)}{n} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0,$$

which allows us to apply the strip criterion.  $\square$

## References

- [Anc88] A. Ancona, *Positive harmonic functions and hyperbolicity*, Potential theory – surveys and problems (Prague, 1987), Lecture Notes in Mathematics, vol. 1344, Springer-Verlag, Berlin, 1988, pp. 1–23.
- [Ave74] A. Avez, *Théorème de Choquet–Deny pour les groupes à croissance non exponentielle*, C. R. Acad. Sci. Paris Sér. A **279** (1974), 25–28.
- [Bea83] A. F. Beardon, *The geometry of discrete groups*, Graduate Texts in Mathematics, vol. 91, Springer-Verlag, New York, 1983.
- [Bil99] P. Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999.
- [Bla55] D. Blackwell, *On transient Markov processes with a countable number of states and stationary transition probabilities*, Ann. Math. Statist. **26** (1955), 654–658.
- [Bri63] J. L. Britton, *The word problem*, Ann. of Math. (2) **77** (1963), 16–32.
- [Bro03] S. Brofferio, *How a centred random walk on the affine group goes to infinity*, Ann. Inst. H. Poincaré Probab. Statist. **39** (2003), no. 3, 371–384.
- [Bro06] ———, *The Poisson boundary of random rational affinities*, Ann. Inst. Fourier (Grenoble) **56** (2006), no. 2, 499–515.
- [BS62] G. Baumslag and D. Solitar, *Some two-generator one-relator non-Hopfian groups*, Bull. Amer. Math. Soc. **68** (1962), 199–201.
- [BSSW15] A. Bendikov, L. Saloff-Coste, M. Salvatori, and W. Woess, *Brownian motion on treebolic space: escape to infinity*, Rev. Mat. Iberoam. **31** (2015), no. 3, 935–976.
- [CD60] G. Choquet and J. Deny, *Sur l'équation de convolution  $\mu = \mu * \sigma$* , C. R. Acad. Sci. Paris **250** (1960), 799–801.
- [CF51] K. L. Chung and W. Fuchs, *On the distribution of values of sums of random variables*, Mem. Amer. Math. Soc. **1951** (1951), no. 6, 12.
- [CFM04] B. Cook, E. M. Freden, and A. McCann, *A simple proof of a theorem of Whyte*, Geom. Dedicata **108** (2004), 153–162.
- [CK12] V. Climenhaga and A. Katok, *Measure theory through dynamical eyes*, 2012, [arXiv:1208.4550](https://arxiv.org/abs/1208.4550).
- [CKW94] D. I. Cartwright, V. A. Kaimanovich, and W. Woess, *Random walks on the affine group of local fields and of homogeneous trees*, Ann. Inst. Fourier (Grenoble) **44** (1994), no. 4, 1243–1288.
- [CW92] D. I. Cartwright and W. Woess, *Infinite graphs with nonconstant Dirichlet finite harmonic functions*, SIAM J. Discrete Math. **5** (1992), no. 3, 380–385.
- [Éli82] L. Élie, *Comportement asymptotique du noyau potentiel sur les groupes de Lie*, Ann. Sci. École Norm. Sup. (4) **15** (1982), no. 2, 257–364.
- [Ers04] A. Erschler, *Boundary behavior for groups of subexponential growth*, Ann. of Math. (2) **160** (2004), no. 3, 1183–1210.
- [Ers10] ———, *Poisson–Furstenberg boundaries, large-scale geometry and growth of groups*, Proceedings of the International Congress of Mathematicians, vol. II, Hindustan Book Agency, New Delhi, 2010, pp. 681–704.
- [Eym72] P. Eymard, *Moyennes invariantes et représentations unitaires*, Lecture Notes in Mathematics, vol. 300, Springer-Verlag, Berlin–New York, 1972.

## REFERENCES

---

- [Fel71] W. Feller, *An introduction to probability theory and its applications*, vol. II, second ed., John Wiley & Sons, Inc., New York–London–Sydney, 1971.
- [FM98] B. Farb and L. Mosher, *A rigidity theorem for the solvable Baumslag-Solitar groups*, *Invent. Math.* **131** (1998), no. 2, 419–451, with an appendix by Daryl Cooper.
- [Fur63] H. Furstenberg, *A Poisson formula for semi-simple Lie groups*, *Ann. of Math. (2)* **77** (1963), 335–386.
- [Fur71] ———, *Random walks and discrete subgroups of Lie groups*, *Advances in Probability and Related Topics*, vol. 1, Dekker, New York, 1971, pp. 1–63.
- [Fur73] ———, *Boundary theory and stochastic processes on homogeneous spaces*, *Harmonic analysis on homogeneous spaces (Proc. Sympos. Pure Math., vol. XXVI, Williams Coll., Williamstown, Mass., 1972)*, Amer. Math. Soc., Providence, R.I., 1973, pp. 193–229.
- [Hae73] J. Haezendonck, *Abstract Lebesgue–Rohlin spaces*, *Bull. Soc. Math. Belg.* **25** (1973), 243–258.
- [Kai91] V. A. Kaimanovich, *Poisson boundaries of random walks on discrete solvable groups*, *Probability measures on groups X (Oberwolfach, 1990)*, Plenum, New York, 1991, pp. 205–238.
- [Kai96] ———, *Boundaries of invariant Markov operators: the identification problem*, *Ergodic theory of  $\mathbb{Z}^d$  actions (Warwick, 1993–1994)*, London Math. Soc. Lecture Note Ser., vol. 228, Cambridge Univ. Press, Cambridge, 1996, pp. 127–176.
- [Kai00] ———, *The Poisson formula for groups with hyperbolic properties*, *Ann. of Math. (2)* **152** (2000), no. 3, 659–692.
- [KKR04] V. A. Kaimanovich, Y. Kifer, and B.-Z. Rubshtein, *Boundaries and harmonic functions for random walks with random transition probabilities*, *J. Theoret. Probab.* **17** (2004), no. 3, 605–646.
- [Kle14] A. Klenke, *Probability theory*, second ed., Universitext, Springer, London, 2014.
- [KV83] V. A. Kaimanovich and A. M. Vershik, *Random walks on discrete groups: boundary and entropy*, *Ann. Probab.* **11** (1983), no. 3, 457–490.
- [KW02] V. A. Kaimanovich and W. Woess, *Boundary and entropy of space homogeneous Markov chains*, *Ann. Probab.* **30** (2002), no. 1, 323–363.
- [LP15] R. Lyons and Y. Peres, *Poisson Boundaries of Lamplighter Groups: Proof of the Kaimanovich–Vershik Conjecture*, 2015, [arXiv:1508.01845](https://arxiv.org/abs/1508.01845).
- [Mei08] J. Meier, *Groups, graphs and trees*, London Mathematical Society Student Texts, vol. 73, Cambridge University Press, Cambridge, 2008.
- [Mes72] S. Meskin, *Nonresidually finite one-relator groups*, *Trans. Amer. Math. Soc.* **164** (1972), 105–114.
- [Neu54] B. H. Neumann, *An essay on free products of groups with amalgamations*, *Philos. Trans. Roy. Soc. London. Ser. A.* **246** (1954), 503–554.
- [Pól21] G. Pólya, *Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz*, *Math. Ann.* **84** (1921), no. 1-2, 149–160.
- [Roh52] V. A. Rohlin, *On the fundamental ideas of measure theory*, *Amer. Math. Soc. Translation* **1952** (1952), no. 71, 55.
- [Ros81] J. Rosenblatt, *Ergodic and mixing random walks on locally compact groups*, *Math. Ann.* **257** (1981), no. 1, 31–42.
- [Rot95] J. J. Rotman, *An introduction to the theory of groups*, fourth ed., Graduate Texts in Mathematics, vol. 148, Springer-Verlag, New York, 1995.

## REFERENCES

---

- [Rou14] G. G. Roussas, *An introduction to measure-theoretic probability*, second ed., Elsevier/Academic Press, Amsterdam, 2014.
- [Rud90] D. J. Rudolph, *Fundamentals of measurable dynamics*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1990.
- [Sav10] E. Sava, *A note on the Poisson boundary of lamplighter random walks*, Monatsh. Math. **159** (2010), no. 4, 379–396.
- [Sch81] R. Schott, *Irrfahrten auf nicht mittelbaren homogenen Räumen*, Arbeitsberichte des Mathematischen Instituts der Universität Salzburg **1-2** (1981), 63–76.
- [VK79] A. M. Vershik and V. A. Kaimanovich, *Random walks on groups: boundary, entropy, uniform distribution*, Soviet Math. Dokl. **20** (1979), no. 6, 1170–1173.
- [Wil70] S. Willard, *General topology*, Addison–Wesley Publishing Co., Reading, Mass.–London–Don Mills, Ont., 1970.
- [Woe89] W. Woess, *Boundaries of random walks on graphs and groups with infinitely many ends*, Israel J. Math. **68** (1989), no. 3, 271–301.
- [Woe00] ———, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, Cambridge, 2000.
- [Woe09] ———, *Denumerable Markov chains*, EMS Textbooks in Mathematics, European Mathematical Society (EMS), Zürich, 2009.